



## International Journal of Social Economics

Inequality in the Distribution of Incomes: A Microstatistical Approach Borrowed from Statistical Thermodynamics

Gundolf H. Kohlmaier Matthias K.B. Lüdeke

### Article information:

To cite this document:

Gundolf H. Kohlmaier Matthias K.B. Lüdeke, (1993), "Inequality in the Distribution of Incomes", International Journal of Social Economics, Vol. 20 Iss 10 pp. 23 - 45

Permanent link to this document:

<http://dx.doi.org/10.1108/03068299310045389>

Downloaded on: 19 February 2016, At: 05:17 (PT)

References: this document contains references to 0 other documents.

To copy this document: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

The fulltext of this document has been downloaded 100 times since 2006\*

Access to this document was granted through an Emerald subscription provided by emerald-srm:514662 []

### For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit [www.emeraldinsight.com/authors](http://www.emeraldinsight.com/authors) for more information.

### About Emerald [www.emeraldinsight.com](http://www.emeraldinsight.com)

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

\*Related content and download information correct at time of download.

---

# Inequality in the Distribution of Incomes

Inequality in  
Distribution of  
Incomes

## A Microstatistical Approach Borrowed from Statistical Thermodynamics

23

Gundolf H. Kohlmaier and Matthias K.B. Lüdeke  
*J.W. Goethe Universität, Frankfurt am Main, Germany*

The following article is an interdisciplinary attempt to bridge the gap between natural and socio-economic sciences. As the authors have their home base in physical chemistry and physics, the approach to economics is unorthodox.

Wilfredo Pareto, the great economist who lived around the turn of the last century, is mostly known for his contributions to welfare economics. The classical goal of welfare economics is to find a measure of potential alternative economic states and select from them the optimum, both with respect to the optimal factor combination in the production function and the optimal product combination in the utility function of the consumer (e.g. see Henderson and Quandt, 1973).

Pareto, however, is less well known for his work on the measured distribution of incomes and wealth among a given population, published by himself under the title, *Courbe de la répartition de la richesse*. The reason for this is perhaps due to the fact that no theoretical framework was established to explain, with *ab initio* principles, the profoundly unequal distributions of income and wealth which are to be found. Such textbooks as that by Samuelson and Nordhaus (1990), describe the issues surrounding the distribution of incomes as still the most controversial in all economics. A comparison of the abilities of individuals, with corresponding distributions of incomes, shows that abilities are much more equally distributed than income. Differences in education, in the chosen occupation or in social circumstances are often made responsible for unequal distribution, although the role of chance and luck affecting the individual's economic circumstances are not excluded (Samuelson and Nordhaus, 1990, p. 651).

We are particularly interested here in the process of the distribution of incomes which, through regular paychecks, is repeated within a chosen time

---

This article is based on a paper presented at the First International Meeting on Ecological Economics in Washington, DC, 1990. The paper attempts to point out a principal approach to evaluate actually occurring distributions in economics and is thought to be more than a formal analogy to theoretical physics and chemistry.

We thank Christof Hager for scientific support, Stefan Habermehl and Andreas Nadler for technical help and Dr R. Hauser and K. Müller for providing statistics of incomes of the Federal Republic of Germany, East and West.

International Journal of Social  
Economics, Vol 20 No 10, 1993,  
pp 23-45 © MCB University  
Press, 0306-8293

interval – i.e. every month or any other time span – which is short compared with the lifetime of an individual. For convenience, we keep the sum of all incomes constant in time. We want to derive an expression for the division of national income, here simply the sum of all incomes, for which all households within a given population compete. We will consider, in particular, two limiting statistical cases. In *Case I*, all households have identical characteristics and chances to participate in the next round of income distribution; in *Case II*, previous income history matters. We distinguish here two sub-cases: in *Case II(a)*, households with previous higher incomes have a better chance to receive high incomes in the future as well; in *Case II(b)*, households with previous differences in income tend to become more equal as time passes.

Counter-intuitive to expectations, a *Case I* statistical distribution of national income among a large number of households does not lead to an equal distribution for all participants, but rather to a skewed distribution, where the lower percentile of the population shares an under-proportional fraction of the income, as is typical in nearly all Lorenz curves.

Our model of *Case I* has a restricted analogy to the statistical physics of a large collection of identical molecules (households) which share a limited amount of total energy (total income), which is distributed among molecules in a number of energy states (income classes). The statistical weight  $\Omega$  (formerly described by the symbol  $W$  from the German word *Wahrscheinlichkeit*) of such a distribution has been characterized by Boltzmann as:

$$\Omega = \frac{N!}{n_1!n_2!\dots n_k!} \quad (1)$$

where  $N$  refers to the total number of molecules (households) and  $n_i$  to the number of molecules (households) in a given energy state (income class).  $\Omega$  characterizes a certain macrostate, in as much as it counts all distinguishable microstates which can be enumerated, as will be shown below, by conventional labelling, i.e. numbering the individual entities.

Out of a large number of macrostates with different weights  $\Omega$ , the one which leads to a maximum  $\Omega$  is selected, under restricting conditions, so that the total number of particles and total energy is conserved.

It is this quantity,  $\Omega_{\max}$ , that Boltzmann could relate to the thermodynamically-derived state function of entropy,  $S$ :

$$S = k_B \ln \Omega_{\max} \quad (2)$$

where  $k_B$  is the Boltzmann constant in statistical physics.

We shall find it convenient to determine  $\Omega$  and  $\Omega_{\max}$  for economical problems as well; for purposes of distinction from the Boltzmann constant, we shall introduce an, as yet, undetermined constant  $k_S$  ( $S$  standing for socio-economics) of dimensions of entropy. If, by external force, another equilibrium state can be maintained with  $\Omega < \Omega_{\max}$  and we remove the force, then the system will have a spontaneous tendency to go into the system with the higher statistical weight.

Identifying the final state with  $S_2 (\Omega_{\max})$ , and the initial state  $S_1 (\Omega)$ , we obtain the famous second law of thermodynamics which suggests that, for an isolated system, any spontaneous process will lead to an increase in entropy:

$$S_2 - S_1 = \Delta S > 0. \quad (3)$$

One of the main points to be made in the following article then, is that the income distribution within a population may have a statistical limit, which corresponds to an extremely unjust social situation. To obtain a socially more acceptable situation, several external societal forces may be necessary to maintain a more just distribution.

The concept of entropy in the analysis of economic problems was first introduced by Georgescu-Roegen, in his well-known book, *The Entropy Law and the Economic Process* (1971). Although the basic roots of entropy within the framework of physics and philosophy, and its possible usefulness in economics, are covered very comprehensively here, very few practical applications for the evaluation of economic processes are given. Only much later do Faber *et al.* (1987) in their monograph, *Entropy, Environment and Resources*, address the economically important question of the environment as a source of the capital required for production, as well as a sink for waste products and its role in recycling of materials.

There are still very great difficulties, however, in applying thermodynamics to economics other than to problems which relate to the production or consumption processes themselves. Here, analogies can be found to the processes which relate to physical or chemical systems which interchange, often through cleverly designed machines which exchange heat, work and material with the environment. For all other processes it seems that the entropy concept is introduced, *a posteriori*, in economic functions of interest. By contrast, we here try to derive distribution functions and relate these directly to entropy, using the microstatistical approach, without ever referring to the very difficult transformation of classical thermodynamic functions into corresponding economic functions.

### Setting the Stage: Representations of the Distribution in Incomes in Economics

The most common representation of the distribution of incomes in a given population is the Lorenz graph where the cumulative fraction of incomes, beginning at the lowest end, is plotted against the cumulative fraction of households receiving these incomes. The Lorenz curve can always be derived from the distribution function  $F(\epsilon; a_1, a_2, \dots)$  of the income  $\epsilon$  ( $\epsilon$  representing the first letter of the German word *Einkommen*, at the same time in physics representing the first letter of energy) and a given set of parameters  $a_i$ .

Let:

$$F(\epsilon, a_1, a_2, \dots) \quad (4a)$$

be the distribution function of income  $\epsilon$  and let:

$$dN = F(\epsilon)d\epsilon \quad (4b)$$

be the number of households in the income bracket between  $\epsilon$  and  $\epsilon + d\epsilon$ . The normalization condition yields that the total number of households,  $N_T$ , is given by the integral from zero to infinity:

$$N_T = \int_0^{\infty} F(\epsilon)d\epsilon \quad (5a)$$

while the number of households below and above  $\epsilon$ :

$$N_{\leq\epsilon} = \int_0^{\epsilon} F(\epsilon)d\epsilon \quad (5b)$$

$$N_{>\epsilon} = \int_{\epsilon+}^{\infty} F(\epsilon)d\epsilon \quad (5c)$$

is given by the above expressions. Similarly, total income is given by the integral:

$$E_T = \int_0^{\infty} \epsilon F(\epsilon)d\epsilon \quad (6a)$$

while the total income below and equal to  $\epsilon$ :

$$E_{\leq\epsilon} = \int_0^{\epsilon} \epsilon F(\epsilon)d\epsilon \quad (6b)$$

and the corresponding total income above  $\epsilon$ :

$$E_{>\epsilon} = \int_{\epsilon+}^{\infty} \epsilon F(\epsilon)d\epsilon \quad (6c)$$

is given by the corresponding expressions. It follows from the definition that the average income of a household is given by:

$$\bar{\epsilon} = \frac{E_T}{N_T}.$$

In the Lorenz graph the cumulative fraction of income below and equal to  $\epsilon$  is plotted versus the cumulative fraction of households below and equal to  $\epsilon$ :

$$\frac{E_{\leq\epsilon}}{E} \text{ vs. } \frac{N_{\leq\epsilon}}{N_T}. \quad (7)$$

As fractions appear in the income, the Lorenz expression is, of course, independent of the units used for income.

For some purposes, illustrated below, it is more convenient to introduce:

$$f(\varepsilon) = \frac{F(\varepsilon)}{N_T} \tag{8}$$

This is a distribution function which is normalized to unity. The base case, which will be discussed in detail throughout this article, is the normalized exponential distribution function:

$$f(\varepsilon; \beta) = \beta e^{-\beta\varepsilon} \quad \beta > 0 \tag{9}$$

where the parameter  $\beta$  characterizes the steepness of the exponential fall off. Within the Lorenz graph representation we can derive:

$$\hat{n}(\varepsilon) \equiv \frac{N_{\leq \varepsilon}}{N_T} = 1 - e^{-\beta\varepsilon} \quad 0 \leq n \leq 1 \tag{10}$$

and

$$\hat{e}(\varepsilon) \equiv \frac{E_{\leq \varepsilon}}{E_T} = 1 - (1 + \beta\varepsilon)e^{-\beta\varepsilon} \quad 0 \leq e \leq 1. \tag{11}$$

As both  $\varepsilon$  and  $\beta$  are positive,  $\beta\varepsilon$  is always positive, and  $\hat{e} \leq \hat{n}$ , which is characteristic for any Lorenz curve. We also notice that the variable  $\varepsilon$  is always connected with  $\beta, \beta\varepsilon \equiv x$  and can be thought of as a dimensionless reduced variable of income:

$$\bar{\varepsilon} = \beta \int_0^{\infty} \varepsilon e^{-\beta\varepsilon} d\varepsilon = \frac{1}{\beta} \tag{12}$$

$x$  represents  $\varepsilon/\bar{\varepsilon}$ . From equation (10) we can determine the inverse function, namely  $\hat{\varepsilon}(\hat{n})$ , which we introduce in equation (11) to obtain:

$$\hat{e}(\hat{n}) = \hat{n} + (1 - \hat{n}) \ln(1 - \hat{n}) \tag{13}$$

where  $\hat{e}(\hat{n})$  is now independent of the parameter  $\beta$ , which was eliminated in the transformation and which depended only on the product  $\beta\varepsilon$ . The two limiting cases  $\hat{n} = 0$ , which yields  $\hat{e}(0) = 0$ , and  $\hat{n} = 1$  which yields  $\hat{e}(1) = 1$ , are correctly represented by equation (13) as:

$$\lim_{\hat{n} \rightarrow 1} \ln(1 - \hat{n})^{(1 - \hat{n})} = 0.$$

We now derive the Gini coefficient for the exponential distribution:

$$C_{\text{Gini}}^{\text{exp}} = 2 \left( \frac{1}{2} - \int_0^1 [\hat{n} + (1 - \hat{n}) \ln(1 - \hat{n})] d\hat{n} \right) \tag{14}$$

$$C_{\text{Gini}}^{\text{exp}} = \frac{1}{2}.$$

This is a very interesting case, in as much as the exponential distribution apparently represents the watershed of all other distributions, extremes of which are a Gini coefficient of 0, for an identical income of all households, and a Gini coefficient of 1, for the extreme unequal distribution where the total income is concentrated in a few rich families. We note that the Gini coefficient of 0.5 for the exponential distribution is independent of the parameter  $\beta$ . An increase in the mean income,  $\bar{\varepsilon}$ , remains unnoticed in the Lorenz graph representation. We suggest representing all other distributions, structurally different from the exponential distribution, by a generalized Maxwell-Boltzmann distribution:

$$f_n(\varepsilon; \beta, n) = C_n \varepsilon^{(n/2-1)} e^{-\beta\varepsilon} \quad (15)$$

where, in statistical physics,  $f_n$  is a distribution function of energy  $\varepsilon$  among molecules with  $n$  degrees of freedom, in which  $n$  are integer numbers equal to or larger than one (see Appendix 1). As defined above, we shall here interpret  $\varepsilon$  as income, and  $n$  as a general parameter, which can be any real number above zero as then the corresponding integrals appearing in the Lorenz graphs are finite. The special case of the exponential function is represented by  $n = 2$ . The normalization constant  $C_n$  can be obtained from the definition of the Gamma function  $\Gamma$  in range of  $n > 0$ :

$$\Gamma(n/2) \equiv \int_0^{\infty} x^{(n/2-1)} e^{-x} dx = (n/2 - 1)! \quad (16a)$$

and the integration of  $f_n$ , within the limits of zero and infinity, and the substitution  $x = \beta\varepsilon$ :

$$C_n = \frac{\beta^{n/2}}{(n/2 - 1)!} \quad (16b)$$

We note in passing:

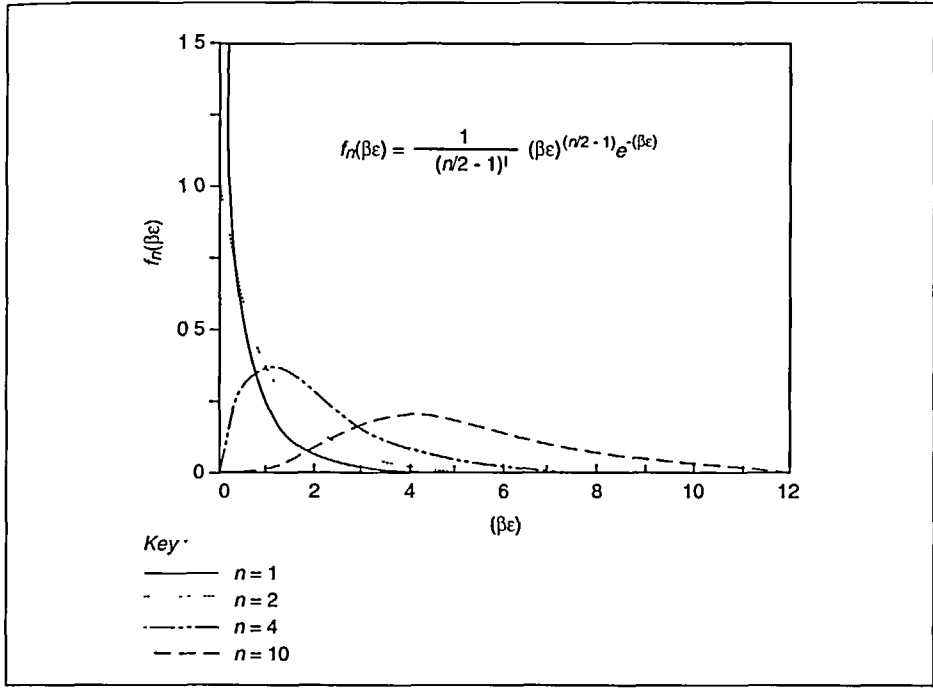
$$\Gamma(n/2) = \frac{\Gamma(n/2 + 1)}{(n/2)} \quad \text{and} \quad \Gamma(1/2) = \sqrt{\pi}.$$

The mean income  $\bar{\varepsilon}$  for the Maxwell-Boltzmann distribution  $f_n(\varepsilon; \beta, n)$ , in  $n$  degrees of freedom, is given by:

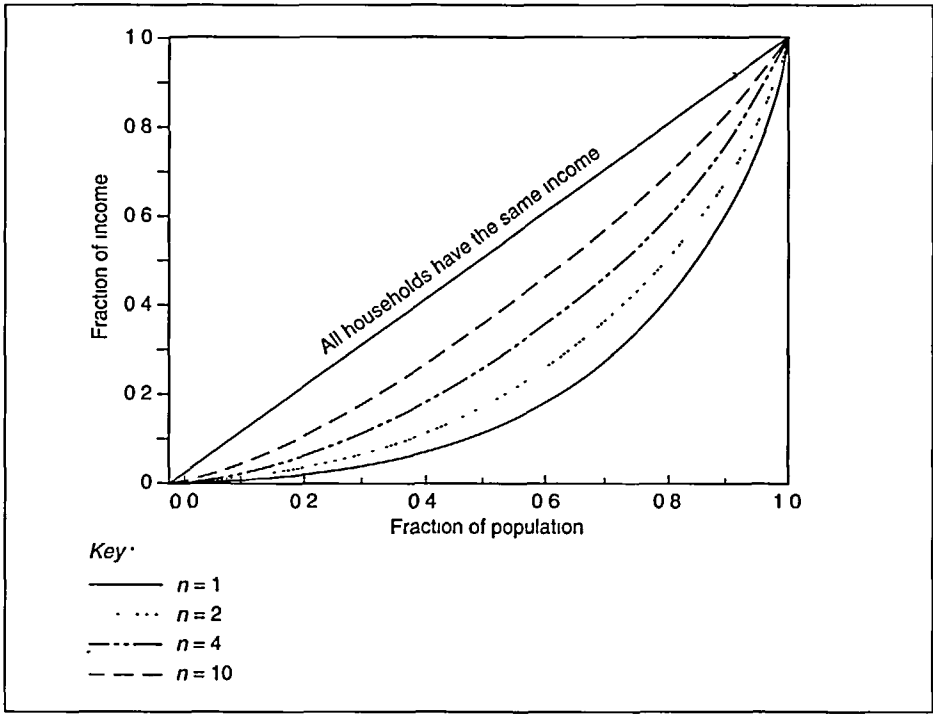
$$\bar{\varepsilon} = \frac{n}{2} \cdot \frac{1}{\beta} \quad (17)$$

which reduces to equation (12) for the special case of an exponential function ( $n = 2$ ). In Figure 1 we present the Maxwell-Boltzmann plot as a function of the (dimensionless) reduced income  $\beta\varepsilon$ , which is particularly convenient as it is independent of the average income  $\bar{\varepsilon}$ . In Figure 2 we show the corresponding Lorenz plot, which again is only dependent on the parameter  $n$ , but not on  $\bar{\varepsilon}$ .

# Inequality in Distribution of Incomes



**Figure 1.** Maxwell-Boltzmann Plot for the Distribution of Income versus the Reduced Income,  $\beta\epsilon$ , as a Function of the Parameter  $n$  (Degrees of Freedom)



**Figure 2.** Lorenz Plot for the Fraction of Incomes versus the Fraction of Population as a Function of the Parameters  $n$  (Degrees of Freedom)



We note that, with increasing  $n$ , the Lorenz curve is increasingly approaching the line where all households have the same income.

### Determining the Maximum Statistical Weight

It is a common method in physics, as well as in economics, to determine the maximum (or minimum) of a function of several variables under some restrictive boundary conditions by Lagrange's method. Here, we seek the maximum of  $\Omega$ , or  $\ln \Omega$ , respectively, because it represents the most likely macrostate in an equilibrium situation. Let  $n_i$  be the number of households in the income class  $i$  and let the macro-distribution be characterized by the statistical weight  $\Omega$  of equation (1). With no restriction other than that the total number of households  $N$  within a population is constant, we introduce the implicit function:

$$F = N - \sum_{i=1}^K n_i = 0 \quad (18)$$

as the only boundary condition. The corresponding Lagrange function is then constructed by adding to  $\ln \Omega$  the function  $F$ , multiplied by the yet undetermined parameter  $\lambda$ :

$$L = \ln \Omega + \lambda F. \quad (19)$$

$L$  is then maximized, with respect to the number of households in each income class  $n_i$ :

$$\frac{\partial L}{\partial n_i} = L_{n_i} = \frac{\partial \ln \Omega}{\partial n_i} + \lambda \frac{\partial F}{\partial n_i} = 0 \quad i = 1, 2, \dots, K \quad (20a)$$

and with respect to  $\lambda$ :

$$\frac{\partial L}{\partial \lambda} = L_\lambda = F = 0 \quad (20b)$$

which gives back the original conditions of equation (18).

The derivative of  $\partial \ln \Omega / \partial n_i$  can be evaluated in simple form if we consider, as in physics, Stirling's first order approximation:

$$\ln n! \approx n \ln n - n \quad (21a)$$

which is valid for larger  $n$ . For smaller  $n$  we need to consider the second-order approximation:

$$\ln n! \approx (n + 0.5) \ln n - n + 0.5 \ln(2\pi) \quad (21b)$$

which gives excellent results even down to  $n = 2$ .

From equation (20a) we obtain with the first-order approximation:

$$n_i = e^{-\lambda+1} = \text{const.}$$

which, along with conditions of equations (20b) or (3), respectively, leads to:

$$n_i = \frac{N}{K} \quad \text{for } K \text{ income classes.} \quad (22)$$

This implies an equal probability for all income classes. If, however, we introduce the second boundary condition, that the total income  $E$  of all households is restricted to a finite quantity and is constant, we need to consider an extended Lagrange function:

$$L = \ln \Omega + \lambda F + \mu G \quad (23)$$

where  $G$  is now the implicit function for income restriction:

$$G = E - \sum_{i=1}^K n_i \varepsilon_i \quad (24)$$

where  $\varepsilon_i$  is now the average income of class  $i$ .

If, again, this new Lagrangian function is maximized we obtain the well-known result of statistical thermodynamics, the Boltzmann distribution:

$$n_i = \frac{N}{z} e^{-\beta \varepsilon_i} \quad (25a)$$

where  $z$  is here the sum of states (*Zustandssumme*) up to level  $K$ :

$$z = \sum_{i=1}^K n_i = \sum_{i=1}^K e^{-\beta \varepsilon_i} \quad (25b)$$

The parameter  $\beta$  determines the fall-off of the exponential;  $e^{-\beta}$  gives the ratio of the number of households in two adjacent income classes:

$$n_{i+1} / n_i = e^{-\beta} \quad (25c)$$

for  $\varepsilon_{i+1} - \varepsilon_i = 1$ .

The parameter  $\beta$  is obtained from the mean income  $\bar{\varepsilon}$  over all income classes:

$$\bar{\varepsilon} = \frac{E}{N} = \frac{\sum_{i=1}^K \varepsilon_i n_i}{\sum_{i=1}^K n_i} = \frac{\sum_{i=1}^K \varepsilon_i e^{-\beta \varepsilon_i}}{\sum_{i=1}^K e^{-\beta \varepsilon_i}} \quad (26)$$

A note of caution should be added here; for a limited and homogeneously spaced number of income classes, the parameter  $\beta$  can, in principle at least, become zero or negative. In fact, depending on the average income  $\bar{\varepsilon}$ , we can distinguish three cases:

- (1)  $\beta > 0$  exponentially decreasing function.
- (2)  $\beta = 0$  number of households in all income classes are identical.
- (3)  $\beta < 0$  exponentially increasing function.

If we consider, for example, a system of five income classes with equal spacings and income levels of  $\varepsilon_1 = \varepsilon_0$ ,  $\varepsilon_2 = \varepsilon_0 + 1$ ,  $\varepsilon_3 = \varepsilon_0 + 2$ ,  $\varepsilon_4 = \varepsilon_0 + 3$ ,  $\varepsilon_5 = \varepsilon_0 + 4$  where  $\varepsilon_0$  is a minimum income level of, for example,  $\varepsilon_0 = 0.5$  (all in arbitrary units of income, e.g. 20,000 DM/year) then  $e^{-\beta} \equiv x$  can be derived from equation (26):

$$\bar{\varepsilon} = \frac{\varepsilon_0 + (1 + \varepsilon_0)x + (2 + \varepsilon_0)x^2 + (3 + \varepsilon_0)x^3 + (4 + \varepsilon_0)x^4}{1 + x + x^2 + x^3 + x^4} \quad (27)$$

$$= \varepsilon_0 + \frac{x + 2x^2 + 3x^3 + 4x^4}{1 + x + x^2 + x^3 + x^4}.$$

By prescribing the average income  $\bar{\varepsilon}$ ,  $x$  and therefore  $\beta$  is determined. The boundary case  $\beta = 0$ , and with it  $x = 1$ , leads to an average income of  $\bar{\varepsilon} = 2.5$  as can be seen easily from equation (27).

Now, in economics, as well as in physics, the number of income classes (energy states) has no upper limit, or has an upper limit which is very large relative to the spacings of the lower classes. This implies that the total income (total energy) for the entire population stays only finite for an exponentially decreasing function with increasing levels of income. The same is done in physics, where Boltzmann found that

$$\beta = \frac{1}{k_B T},$$

where  $k_B$  is the Boltzmann constant and  $T$  is the absolute temperature (temperature in Celsius + 273.15). Only for  $T \rightarrow \infty$  does  $\beta = 0$  act as a limiting case, but never negative.

We can ask here, also, whether the function  $\Omega$ , or  $\ln \Omega$  has a minimum. This would be a state of distribution which would be extremely unlikely, at least for a closed system of identical entities.

A state of low  $\Omega$ , and therefore low  $S$ , can only be maintained in an open system where it is possible to import negentropy or to export entropy. Living systems are always such open systems, and human societies are good representatives. Households within a population are "rather smart" in the competing economic market; an ensemble of molecules is *per se* "rather stupid"; it depends on the sophistication of machines, which exchange energy and materials with their surroundings, as to whether they are able to be constructive and maintain a low entropy.

An examination of  $\Omega$  shows that there is no regular minimum in the sense that the partial derivatives are zero. But there are  $K$  boundary minima, namely when:

$$n_i = 0 \text{ for all } i \neq j$$

$$n_i = N \text{ for } i = j$$

then

$$\Omega = \frac{N!}{0! 0! 0! \dots N! \dots 0!} = 1. \quad (28)$$

Which  $n_i$  becomes  $N$  depends on the boundary condition of the total income  $E$  for:

$$\sum_{j=0}^K n_j \varepsilon_j = E.$$

If we translate this state into the corresponding entropy function, we obtain:

$$S_{\min} = k_S \ln \Omega_{\min} = 0 \quad (29)$$

where  $\Omega_{\min}$  is a selective situation, which can be maintained in a steady state only through the smartness of the individual households or through active interference by some governmental restrictions.

In the same light of argument, a bimodal distribution where some of the economic entities dominate the mass of the remaining economic entities, will be a state of still low entropy compared to the statistical result. It will be higher, however, than the minimum situation described above.

### Limiting Cases for the Distribution Problem as Derived from the Microstatistical Method

We are interested, here, in how total products derived from economic activities are divided among households, disregarding any other groups which participate in the market. We refer here to one part of the factor market and study income as derived from the sale of services of the mind and body of family members and from the sale of services of properties which the household owns. We will first start with what Boulding economic analysis, called the *principle of equal advantage* (Boulding, 1966):

If the owners of any resources think these can be put to better advantage in some other use than the one in which they are employed, the resources will be transferred from the less advantageous to the more advantageous use. The process of transfer will generally have the effect of making the occupation into which resources have been transferred less advantageous than before; it will make the occupation out of which resources, have been transferred more advantageous than before. As long, therefore, as there are people who believe that the resources they own, be they their own bodies or some other object, will yield them a greater advantage in an occupation different from that in which they are at present employed, then resources will be transferred from one occupation to another

This results in a market economy, with no restriction to migrations where, an equilibrium is achieved in which the (monetary and non-monetary) advantage is the same in all occupations. This limiting case corresponds to the minimum entropy case derived above, in which all households have the same income. Of course, we know that this is not true in reality. We therefore now explain the other limiting case of *equal probability of all microstates*. This implies that all

households compete for the product of economic activity in purely random fashion.

The evaluation of different macrodistributions and their statistical weight, represented by  $\Omega$  (equation 1), will be discussed now. We choose  $N$ , the number of households, and  $K$ , the number of income classes which are small and restricted with an upper limit. For illustration we discuss the possible distributions of  $N = 10$  households among  $K = 5$  income classes as above:  $\varepsilon_1 = 1/2$ ,  $\varepsilon_2 = 3/2$ ,  $\varepsilon_3 = 5/2$ ,  $\varepsilon_4 = 7/2$ ,  $\varepsilon_5 = 9/2$ ; all in arbitrary units of income, e.g. 20,000 DM/year. We shall discuss in detail three cases, namely one case where the total income is restricted to 15 income units, representing the exponential case of an average income  $\varepsilon = 1.5$ ; a second case where the total income is 25 income units, representing the case of equal distribution among all income classes ( $\beta = 0$ ) with an average income of  $\varepsilon = 2.5$ ; and a third case where the total income is 35 units, corresponding to the inverted exponential distribution ( $\beta < 0$ ) with  $\varepsilon = 3.5$ , where all cases refer to five equally spaced income classes. It should be noted here that the relative weight of the statistical distributions does not change (as is evident from equation (27)) whether we consider ten households, or ten household units, where one unit represents 100,000 people, one or ten million people.

However, the statistic becomes very opaque because, already, without a total income restriction and equal a priori probability for each of the income classes, we need to distinguish

$$K^N = 5^{10} = 9,765,625$$

microdistributions, if we attached to each a family a number out of the set  $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ .

We first discuss the case for  $\varepsilon = 1.5$ , and describe the results in Table I. We recognize 23 different macrostates among 72,403 microdistributions. Each macrostate is characterized by a characteristic income distribution, the different possibilities of realization can be derived from  $\Omega$ , given in equation (1). Macrodistribution XXIII has a statistical weight of one, and is the most unlikely state – equal income of all households is the least favourable in a statistical sense. On the other hand, macrodistribution I, which is close to an exponential distribution, has the highest statistical weight, with the largest fraction of the population in the lowest state. The exact exponential distribution cannot be realized because of the small number of households participating in the distribution.

We now discuss the case for  $\varepsilon = 2.5$ , as described in Table II. Average energy has been chosen according to equation (27) such that  $e^{-\beta} = 1$  and therefore  $\beta = 0$ . Because of the large number of macrodistributions possible, namely 55, which are not too interesting in detail, we have listed only the distributions with high and low statistical weight. We notice here that the distribution with an equal number of households in each income class, represented by state I, has the highest statistical weight. Although not immediately obvious, it again presents an exponential distribution for  $\beta = 0$ . Macrostate LV with all households

$\epsilon_i$	0.50	1.50	2.50	3.50	4.50	$\Omega$	$\Omega/\Sigma$
I	4	3	2	1	0	12,600	0.1740
II	5	2	1	2	0	7,560	0.1044
III	5	2	2	0	1	7,560	0.1044
IV	4	4	1	0	1	6,300	0.0870
V	3	5	1	1	0	5,040	0.0696
VI	5	1	3	1	0	5,040	0.0696
VII	5	3	0	1	1	5,040	0.0696
VIII	6	1	1	1	1	5,040	0.0696
IX	3	4	3	0	0	4,200	0.0580
X	4	2	4	0	0	3,150	0.0435
XI	4	4	0	2	0	3,150	0.0435
XII	2	6	2	0	0	1,260	0.0174
XIII	6	0	2	2	0	1,260	0.0174
XIV	6	2	0	0	2	1,260	0.0174
XV	3	6	0	0	1	840	0.0116
XVI	6	0	3	0	1	840	0.0116
XVII	6	1	0	3	0	840	0.0116
XVIII	2	7	0	1	0	360	0.0050
XIX	7	0	0	2	1	360	0.0050
XX	7	0	1	0	2	360	0.0050
XXI	5	0	5	0	0	252	0.0035
XXII	1	8	1	0	0	90	0.0012
XXIII	0	10	0	0	0	1	0.0000
$\Sigma =$						72,403	1.0000

Inequality in  
Distribution of  
Incomes

35

**Table I.**  
Macrostates (I, II, . . . ,  
XXIII) and Their  
Statistical Weight  $\Omega$  for  
a Ten-household/Five-  
income Class System  
and a Total Energy of  
 $E = 15$  Income Units

earning the same income, this time increased by one unit above the first case, again has the lowest statistical weight.

For completeness, we have also calculated the case in which  $\beta < 0$  (as shown in Table III) for an average income of 3.5 units. In reality, this case would not occur, as mentioned in the previous section, because the income levels have no upper limit. The example is chosen in such a way that Table I and Table III are mirror images of each other. Here again, the next exponential macrostate I has the highest statistical weight, while macrostate XXIII is realized by just one microdistribution.

In summary, we state the the exponential distribution independent of total income, has the highest chance to be realized, if a pure statistical distribution is applicable. The entropy associated with the state of highest probability for the near exponential distribution is simply proportional to  $\ln \Omega_{\max}$ . Any other distribution is a non-equilibrium state, which can be retained by external forces or a state in a thermodynamically open system. If an entropy is associated with

$\epsilon_{i=}$	0 50	1 50	2 50	3 50	4 50	$\Omega$	$\Omega/\Sigma$
I	2	2	2	2	2	113,400	0.1323
II	1	3	2	3	1	50,400	0.0588
III	1	3	3	1	2	50,400	0.0588
IV	2	1	3	3	1	50,400	0.0588
V	2	3	1	1	3	50,400	0.0588
VI	3	1	1	3	2	50,400	0.0588
VII	3	1	2	1	3	50,400	0.0588
VIII	1	2	4	2	1	37,800	0.0441
IX	1	4	1	2	2	37,800	0.0441
X	2	2	1	4	1	37,800	0.0441
XI	2	2	3	0	3	25,200	0.0294
XII	2	3	0	3	2	25,200	0.0294
XIII	2	1	4	1	2	37,800	0.0441
...	...	...	...	...	...	...	...
XLVI	2	2	0	6	0	1,260	0.0015
XLVII	0	6	1	0	3	840	0.0010
XLVIII	3	0	1	6	0	840	0.0010
XLIX	0	2	7	0	1	360	0.0004
L	1	0	7	2	0	360	0.0004
LI	0	5	0	5	0	252	0.0003
LII	5	0	0	0	5	252	0.0003
LIII	0	1	8	1	0	90	0.0001
LIV	1	0	8	0	1	90	0.0001
LV	0	0	10	0	0	1	0.0000
$\Sigma =$						856,945	1.0000

**Table II.**  
Macrostates (I, II, ..., LV) and Their Statistical Weight  $\Omega$  for a Ten-household/Five-income Class System and a Total Energy of  $E = 25$  Income Units

these non-equilibrium distributions, where  $S$  is again proportional to  $\ln W$ , then the entropy change, according to equation (3), is given by:

$$\Delta S = k_s (\ln \Omega_{\max} - \ln \Omega) = k_s \ln \frac{\Omega_{\max}}{\Omega} \quad (30)$$

$\Delta S$  is positive in all cases. The equilibrium entropy associated with the (near) exponential distribution is dependent on the mean income, as long as  $\beta \geq 0$ ,  $S$  increases with increasing income, as a comparison of  $\Omega$  in Tables I and II shows.

### Transition from the Discrete Class Model to the Continuous Boltzmann Exponential Distribution

In the previous section we presented a model of equal income spacing, considering five income classes, ten households and three different values of

$\varepsilon_i =$	0.50	1.50	2.50	3.50	4.50	$\Omega$	$\Omega/\Sigma$
I	0	1	2	3	4	12,600	0.1740
II	0	2	1	2	5	7,560	0.1044
III	1	0	2	2	5	7,560	0.1044
IV	1	0	1	4	4	6,300	0.0870
V	0	1	1	5	3	5,040	0.0696
VI	0	1	3	1	5	5,040	0.0696
VII	1	1	0	3	5	5,040	0.0696
VIII	1	1	1	1	6	5,040	0.0696
IX	0	0	3	4	3	4,200	0.0580
X	0	0	4	2	4	3,150	0.0435
XI	0	2	0	4	4	3,150	0.0435
XII	0	0	2	6	2	1,260	0.0174
XIII	0	2	2	0	6	1,260	0.0174
XIV	2	0	0	2	6	1,260	0.0174
XV	0	3	0	1	6	840	0.0116
XVI	1	0	0	6	3	840	0.0116
XVII	1	0	3	0	6	840	0.0116
XVIII	0	1	0	7	2	360	0.0050
XIX	1	2	0	0	7	360	0.0050
XX	2	0	1	0	7	360	0.0050
XXI	0	0	5	0	5	252	0.0035
XXII	0	0	1	8	1	90	0.0012
XXIII	0	0	0	10	0	1	0.0000
$\Sigma =$						72,403	1.0000

Inequality in  
Distribution of  
Incomes

37

**Table III.**  
Macrostates (I, II, ..., XXIII) and Their Statistical Weight  $\Omega$  for a Ten-household/Five-income Class System and a Total Energy of  $E = 35$  Income Units

average income:  $\varepsilon = 1.5, 2.5, 3.5$ . As in case 1,  $\varepsilon = 1.5$  represents a typical exponential fall-off curve. We will examine here the transformation of this discrete income level system to the corresponding continuous distribution function of exponential character. In fact, the non-normalized exponential function  $C \cdot e^{-\beta\varepsilon}$  may be partitioned into five income columns of width  $\Delta\varepsilon = 1$  and height  $C \cdot e^{-\beta\varepsilon_i}$  ( $i = 1, 2, \dots, 5$ ), in which the areas or the heights of the column exactly represent the corresponding relative weights of the occupation of the discrete system. With equation (27) we can evaluate the steepness parameter  $\beta$  which, for  $\varepsilon = 1.5$  ( $\varepsilon_0 = 0.5$ ), assumes the value  $\beta = 0.565$  and  $C = 6.12$  for ten households. The area of the five columns can be replaced, to a good approximation, by the corresponding area underneath the exponential curve within the limits of  $\varepsilon = 0$  and  $\varepsilon = 5$ . Redetermining the parameter  $\beta$  for the continuous distribution between the limits of  $\varepsilon = 0$  and  $\varepsilon = 5$  yields a somewhat smaller value, namely  $\beta = 0.53$ . If one compares this value with the



corresponding value of  $\beta$ , when the exponential tail is considered, we obtain with the results of equation (12):

$$\bar{\varepsilon} = \int_0^{\infty} \varepsilon \beta e^{-\beta \varepsilon} d\varepsilon = \frac{1}{\beta}$$

$$\text{or } \beta = \frac{1}{\bar{\varepsilon}} = 0.6.$$

Using this parameter  $\beta$  we can determine the fraction of households between  $\varepsilon = 0$  and  $\varepsilon = 5$  according to equation (10):

$$\hat{n}(5) = 1 - \beta e^{-\beta \varepsilon} = 0.976$$

and the fraction of income  $\varepsilon$  between  $\varepsilon = 0$  and  $\varepsilon = 5$  according to equation (11):

$$\hat{\varepsilon}(5) = 1 - (1 + \beta \varepsilon)^{-\beta \varepsilon} = 0.845$$

which shows that, in the exponential tail, 2.4 per cent of households receive 15.5 per cent of income.

### Dynamic Considerations of the Distribution of Income in Relation to the Statistical Analysis

The Boltzmann statistics of an ensemble of molecules which share a given total energy assumes that the energy can be exchanged through binary, ternary and higher order collisions such that every energy configuration in molecule space can be explored. Macroscopic systems are generally characterized by a very large number of molecules, of the order of  $10^{24}$ , and by collision frequencies of the order of nanoseconds. The statistical weight of the exponential distribution is many orders of magnitude larger than any other non-exponential distribution such that the former distribution represents the equilibrium situation. Any non-equilibrium distribution will relax quickly to the corresponding exponential equilibrium distribution.

We shall now analyse the economic process in a number of dynamic experiments which describe the physics of identical particles correctly and which may bear some relation to an idealized economic situation. In the previous section we could demonstrate that the simple ten-household/five-income class system is a good representative of the more general statistical result derived above. As a first step towards a mechanistic interpretation of the underlying dynamic process we performed the following numerical experiment.

We assumed, for all ten households, that any of the income classes can be occupied with equal probability. In practice a "die" with five income states is thrown ten times in a row associating each household 1, 2, ..., 10, with one of the five income states. This procedure was repeated about  $10^7$  times yielding most of the  $5^{10}$  possible microstates which, for the greater part, do not match the boundary condition that their total energy equals 15 (energy Case 1). From the example in the previous section, we know that there are 72,403 out of  $5^{10}$  microstates (or about 1 per cent) which match the prescribed total income,

constituting 23 distinguishable distributions. Now the repetitive occurrence of each of these distributions was counted and divided by the total number of distributions matching the income condition. The obtained relative frequencies, taken on a cumulative basis, were shown to be equal to the calculated probabilities given in the table of the previous section within a precision of  $\pm 0.04$ . This dynamic model reproduces the expected equilibrium, but includes “transient” states which are taken out because they do not fulfil the total income conservation.

In the previous example, we considered essentially the statistical distribution of national income resulting from all economic activities in the production of goods and services. If we follow up the picture of an exchange in pairs, as is important in physics, we realize that such pair exchanges also exist between economic entities. The most common one is, of course, the exchange of real goods between two households in primitive societies. Income of households usually refers to one part of a factor market where monetary income is derived from the sale of services of mind and body and also from the sale of services of property.

In a situation of competition between many individuals with one employer, we can apply the example of the distribution of the “cake” as above. A more restricting consideration, which still keeps the total income constant, is that any chosen pair of individuals can redistribute their income in the sense that the sum of the two incomes remains constant. If this process is repeated many times for all possible pair combinations, we should indeed arrive at the statistical result of the above. We started with an arbitrary distribution, matching the total income boundary condition. Then we randomly chose pairs of economic entities, simulating their exchange by a stochastic redistribution of the income sum of the two households involved. With  $\varepsilon(i)$  being the income of the entity with number  $i$  we calculate for randomly chosen  $i$  and  $j$ :

$$\{\varepsilon(i), \varepsilon(j)\} \Rightarrow \{r [\varepsilon(i) + \varepsilon(j)], (l-r)[\varepsilon(i) + \varepsilon(j)]\} \quad (31)$$

$r$ : uniformly distributed stochastic variable,  $0 \leq r \leq 1$ .

When using the ten-household/five-income class system as an example, we choose  $i, j$  and  $\varepsilon$  from the sets:

$i, j$  from  $\{1, 2, \dots, 10\}$ , and  $\varepsilon$  from  $\{1/2, 3/2, 5/2, 7/2, 9/2\}$ .

We performed the interaction experiments in sets with  $l$  consecutive exchanges, where  $l$  exceeded the number of different pairs to be formed from  $N$  distinguishable households (here  $N(N-1)/2 = 45$ ). The resulting distribution was identified with one of the 23 macrodistributions, and then the next set of  $l$  exchanges, starting with the same initial distribution, was performed. For our example system we obtained that the relative frequency of the resulting distributions equals the calculated values of the statistical results, independent of which initial distribution was chosen. Hence, we established a model of the redistribution of income among the economic entities which reproduces the

statistical result of the previous section and allows an identification of the exchange partners. It should be noted that the initial distribution can influence the final outcome if, within each set, only a small number of interactions are considered. Furthermore, we emphasize that as the number of entities and classes increases, as in any real system, the weight of the exponential or near exponential distribution exceeds all other allowed distributions many times.

Interpreting a binary exchange as an economic interaction, one may assume that the redistribution between the two partners depends on their state before the interaction, implying a Markov process beyond the analysis of the previous section. We tested the following form of asymmetric interactions for our example system:

$$\{\varepsilon(i), \varepsilon(j)\} \Rightarrow \{r'[\varepsilon(i) + \varepsilon(j)], (1 - r')[\varepsilon(i) + \varepsilon(j)]\} \quad (32)$$

with:  $r' = r(\varepsilon(j) + \alpha) / (\varepsilon(j) + \alpha)$

$r$ : uniformly distributed stochastic variable,  $0 \leq r \leq 1$ .

$\alpha$ : parameter, determining the redistribution of household income with respect to the household income before the exchange ( $\alpha \rightarrow \infty$ : previous model).

The relation  $(\varepsilon(j) + \alpha) / (\varepsilon(i) + \alpha)$ , a measure of the asymmetry in the income redistribution, favours the entity with a higher income. It is most sensitive for  $\alpha \rightarrow 0$ ; for  $\alpha \rightarrow \infty$  it reproduces the independence from the income before exchange. Using this model of exchange for arbitrary finite  $\alpha$ , we obtained the result that, independent of the initial distribution of income, a bimodal distribution (see distributions XIX and XX of Table I) is favoured over all other distributions with the relative frequency approaching 1 for  $l \rightarrow \infty$ . This particular case may perhaps be an analogue for the distribution of incomes in underdeveloped nations, where many people with practically no income are contrasted with a few rich people.

We also formulated the principle of equal advantage on the basis of binary exchanges. In an analogy to the above, the following strategy was assigned:

$$\{\varepsilon(i), \varepsilon(j)\} \Rightarrow \{r'[e(i) + e(j)], (1 - r')[e(i) + e(j)]\} \quad (33)$$

with  $r' = \frac{r^A}{2}$ .

In the case of  $A = 1$ , we get back the unbiased statistical result of an exponential distribution.

If  $0 < A < 1$ , we obtain a tendency towards equilibration, in the sense that the economic unity with the lower income moves to a higher advantage position, lowering at the same time the incomes in the high advantage position. Again, if the experiment is performed in a set of  $l$  consecutive interactions, and then repeated in many sets, we obtain the equal income situation of distribution XXIII. We noticed in the section above that these mechanisms assume an

evolutionary strategy in which the entities are free to move within a market economy, and smart enough to sense their advantage.

**Comparison of Lorenz Curves of Income Derived from Empirical Data and from the Generalized Maxwell-Boltzmann Distribution Function**

In the second section we derived the Lorenz function from an exponential distribution function of income and could show that its Gini coefficient is 1/2, independent of the exponential parameter  $\beta$ . The exponential function itself has the highest statistical weight  $W$ , and therefore is the most stable one in a thermodynamic sense. We developed, in Appendix 1, the more general statistical result for a system with  $n$  degrees of freedom, the Maxwell-Boltzmann distribution, of which the exponential distribution is a special case for  $n = 2$ . In this section, we now explore, for a series of countries, whether the Lorenz curve, derived from the Maxwell-Boltzmann distribution, produces a reasonable match with the corresponding empirical data.

As before, in equations (15) and (16b), for the exponential function, we define here for the generalized, normalized Maxwell-Boltzmann function:

$$f(\epsilon; \beta, n) = \frac{\beta^{n/2}}{\Gamma(n/2)} \epsilon^{(n/2-1)} e^{-\beta\epsilon}$$

the cumulative fraction of the population  $\hat{n}(\epsilon)$ , earning equal or less than  $\epsilon$ :

$$\hat{n}(\epsilon) = \frac{N_{\leq\epsilon}}{N_T} = \frac{\int_0^{\epsilon} f(\epsilon; \beta, n) d\epsilon}{\int_0^{\infty} f(\epsilon; \beta, n) d\epsilon} \tag{34}$$

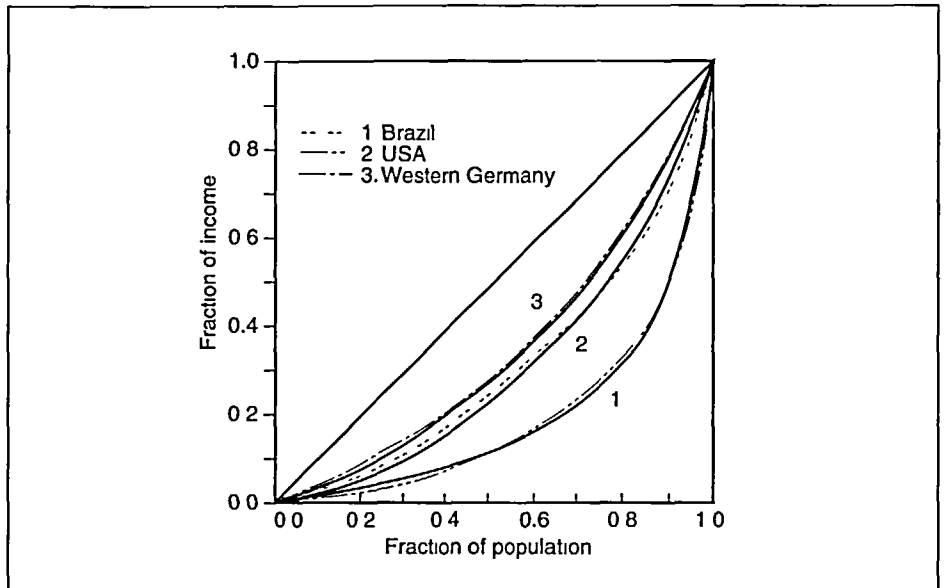
and the cumulative fraction of income earned by households with incomes less than  $\epsilon$ :

$$\hat{e}(\epsilon) = \frac{E_{\leq\epsilon}}{E_T} = \frac{\int_0^{\epsilon} \epsilon f(\epsilon; \beta, n) d\epsilon}{\int_0^{\infty} \epsilon f(\epsilon; \beta, n) d\epsilon} \tag{35}$$

We try now to fit some empirical data of the distribution of income with Lorenz curves derived from the Maxwell-Boltzmann distribution. Some typical results are shown in Figure 3. It can be seen that the Lorenz curves derived from the Maxwell-Boltzmann functions produce a rather good match.

Tables IV and V show the Gini coefficients for the empirical data, the exponential  $n$  for the calculated curves and the mean deviation of the calculated from the empirical curves  $\langle |\hat{e}_{emp} - e_{calc}| \rangle$ .

**Figure 3.**  
Comparison of the  
Empirical and the  
Calculated Lorenz  
Curves for Brazil,  
the USA and Western  
Germany



The results shown in Table IV were obtained by using parameters  $n \geq 0$ . If the distribution of income is very unequal (i.e. larger Gini coefficients) we have to use  $n < 0$ . In these cases, the integrals over  $f(\varepsilon, \beta, n)$  diverge at the lower limit ( $\varepsilon_i = 0$ ) and we need to introduce a low cut-off bound  $\varepsilon_i$  for mathematical reasons to prevent a singularity. Only if  $\varepsilon_i > 0$  does the parameter  $\beta$  become relevant. The result for such a fitting process is given in Table V.

In summary, we can state that the Lorenz curves for the distribution of incomes deviate in general from the generating curve of a simple exponential function, which corresponds to  $n = 2$  in the Maxwell-Boltzmann distribution.

**Table IV.**  
Evaluation of the Gini  
Coefficients for Selected  
Countries within  
Parametrization  $n$  of  
Generalized Maxwell-  
Boltzmann Distribution

Country	$n$	Gini	$\langle  \ell_{\text{emp}} - \ell_{\text{calc}}  \rangle$
FRG-E 90	6.4	0.295	0.002
FRG-W 89	5.9	0.318	0.001
Sweden	6.0	0.301	0.007
UK	5.0	0.327	0.008

**Table V.**  
Evaluation of the Gini  
Coefficient for Brazil

Country	$\varepsilon_i$	$\beta$	$n$	Gini	$\langle  \ell_{\text{emp}} - \ell_{\text{calc}}  \rangle$
Brazil	0.01	0.8	-1.0	0.624	0.009
With a best match with the values $n, \beta$ in the Maxwell-Boltzmann distribution					

The Gini coefficients (which equal 0.5 for this particular case ( $n = 2$ )) either lie above 0.5, as for the developing country Brazil (Gini coefficient = 0.624) or below, as for all of the industrial countries examined here. We note that, for all cases with  $n > 0$ , the results of the Gini coefficient are independent of the parameter  $\beta$ , which is determined by the mean income  $\varepsilon = 0.5 n\beta^{-1}$ . Again, the Lorenz curves are independent of  $\varepsilon$  or  $\beta$ ; thus, an increase in income without structural changes does not show up in the Lorenz curve. We summarize by saying that the income distribution may, at least to some extent, be the result of a statistical distribution process where all allowed microstates have equal statistical weight.

In conclusion, we should like to state that neither of the two extreme standpoints to explain the inequality of incomes – namely, the conventional one of different backgrounds and the one presented here of equilibrium statistics – can lead to a totally satisfactory explanation. We perhaps think that the evolutionary strategies, as developed most recently in molecular biology, may be an interesting lead to follow up in more detail.

#### References and Further Reading

- Boulding, K.E. (1966), *Economic Analysis, Vol I Microeconomics*, p. 720.  
 Faber, M., Niemes, H. and Stephan, G. (1987), *Entropy, Environment and Resources*, Springer Verlag, Berlin, p. 205.  
 Georgescu-Roegen, N (1971), *The Entropy Law and the Economic Process*, Harvard Paper Press, p 457.  
 Henderson, J.M. and Quandt, R E (1973), *Mikroökonomische Theorien*, translated by Meißner, W and Vahlen, F., Munich, p. 393  
 Maxwell, J.C. (1921), *Theory of Heat*, 10th ed, London, p 189.  
 Pareto, V. (1968), *Ecrits sur la Courbe de la Répartition de la Richesse. Travaux de Droit, d'Economie de Sociologie et de Sciences Politiques*, No 36, Librairie Droz, Genève.  
 Samuelson, P.A. and Nordhaus, W.D (1990), *Economics*, 13th ed, McGraw-Hill, NY.

#### Appendix 1. Maxwell-Boltzmann Distribution of Molecules in $n$ Degrees of Freedom – Generalized Maxwell-Boltzmann Distribution of Income

The distribution of velocities in one-, two- and three-dimensional space of an ensemble of molecules, as well as their corresponding kinetic energies, is a well-known result of statistical thermodynamics. Translational, rotational and vibrational energies of molecules can always be expressed as a sum of the squares of the corresponding generalized co-ordinates  $q_i$ :

$$\varepsilon(n) = \sum_{i=1}^n \varepsilon_i = \sum_{i=1}^n \frac{q_i^2}{2} \quad (A1)$$

where the sum extends over  $n$  degrees of freedom, where  $n = 3$  for translation and rotation each, and  $n = 2$  for each vibration as both kinetic and potential energies need to be considered.

The energetic state of the molecule can be represented by a vector  $\vec{r}$  in  $n$ -dimensional space.

$$\vec{r}_{(n)} = (q_1, q_2, \dots, q_n) \quad (A2)$$

where the distance  $r_{(n)} = |r^n|$  from the origin is given by the Pythagorean relation:

$$r_{(n)}^2 = \sum_{i=1}^n q_i^2 \quad (A3)$$

The probability distribution of molecules in two degrees of freedom. (e.g. kinetic and potential energy of a vibrational motion) is given by the well-known Boltzmann exponential function:

$$\frac{dN_{q_1 q_2}}{N} = \exp\left\{-\beta \frac{q_1^2 + q_2^2}{2}\right\} dq_1 dq_2 \quad (A4)$$

which can be related to the more familiar form in energy space, using the relations:

$$\epsilon_{(2)} = \frac{r_{(2)}^2}{2}; d\epsilon_{(2)} = r_{(2)} dr_{(2)}$$

and the transformation of Cartesian co-ordination  $(q_1, q_2)$  into polar co-ordinates  $(r_{(2)}, \varphi)$

$$dq_1 dq_2 = r_{(2)} dr_{(2)} d\varphi$$

resulting after integration over  $\varphi$  in:

$$\frac{dN_{\epsilon_{(2)}}}{N} = 2\pi \exp\{-\beta\epsilon_{(2)}\} d\epsilon_{(2)}. \quad (A5)$$

The general result for the joint probability in  $n$  degrees of freedom is given by a generalized expression of equation (A4)

$$\frac{dN_{q_1, q_2, \dots, q_n}}{N} = \exp\left\{-\beta \sum_{i=1}^n \frac{q_i^2}{2}\right\} dq_1 dq_2 \dots dq_n \quad (A6)$$

or after a transformation to polar co-ordinates in  $n$ -dimensional space and integration over all variables except  $r_{(n)}$ :

$$\frac{dN_{r(n)}}{N} = \text{const. } r_{(n)}^{n-1} \exp\{-\beta r_{(n)}^2 / 2\} dr_{(n)}. \quad (A7)$$

If, again, we make the transition energy space, we obtain the now normalized result of the Maxwell-Boltzmann distribution in  $n$  degrees of freedom:

$$\frac{dN_{\epsilon_{(n)}}}{N} = \frac{1}{\Gamma(n/2)} (\beta\epsilon_{(n)})^{(n/2-1)} \cdot \exp\{-\beta\epsilon_{(n)}\} d(\beta\epsilon_{(n)}) \quad (A8)$$

where  $\Gamma(n/2) = (n/2 - 1)!$

The Maxwell-Boltzmann function evaluates this distribution of energy among molecules according to statistical physics, with the result that the energy distribution is unequal among equal molecules and that it depends on the number of degrees of freedom taken to be equal for all molecules of the set.

#### *A Proposal for a Transition from Statistical Physics to Economics*

Although it is not clear, at present, if there is an economical analogy to energy and the degrees of freedom, we will suppose that we can represent the economic status of individual households in a hyperspace with a dimension given by the degrees of freedom,  $n$ , in which a function of the distance from the origin represents the income. As a first assumption we choose the square of the distance to be representative because, then, the laws of statistical physics are directly applicable.

We shall suppose for the moment that the state variables represent the various possibilities of economic engagement common to all households participating in the economic process. We shall see in the discussion of the article that if the close analogy between statistical physics and the economy is maintained, how the resulting distribution of income depends on the number of the "degrees of freedom". We now interpret  $e_{(n)}$  as the income  $e$  (omitting the index  $n$ ) and use  $n$  as a convenient parameter to find the closest match to the observed Lorenz curves. We will use the normalized function:

$$f_n(\epsilon) = \frac{\beta^n}{\Gamma(n/2)} \epsilon^{(n/2-1)} \exp\{-\beta\epsilon\} \quad (\text{A9})$$

where  $\beta$  is a steepness parameter of the exponential fall off and  $n$  is the parameter for the "degrees of freedom", which gives analytical results in the Lorenz graph interpretation, as long as  $n$  is a real positive number, i.e.  $n > 0$ .

### Appendix 2. The Pareto Function

It is very interesting to note that Pareto suggested a "natural law" for the income distribution of the form:

$$N_{\geq \epsilon} = \frac{A}{(\epsilon + a)^\alpha} e^{-\beta\epsilon} \quad (\text{A10})$$

where  $A$  depends on the size of the population,  $a$  is constant of equal magnitude as the lowest considered income, and where  $\alpha$  ranges between 1.4 and 1.8 with a mean value of 1.5. As:

$$N_{\geq \epsilon} = \int_{\epsilon}^{\infty} F(\epsilon) d\epsilon \Rightarrow \frac{dN_{\geq \epsilon}}{d\epsilon} = -F(\epsilon)$$

we obtain the corresponding distribution function:

$$f(\epsilon) = \frac{A}{N_T} \frac{e^{-\beta\epsilon}}{(\epsilon + a)^\alpha} \left( \beta + \frac{\alpha}{(\epsilon + a)} \right) \quad (\text{A11})$$

This function has great formal similarity to the suggested generalized Maxwell-Boltzmann distribution function, or a sum of two such functions, if we neglect  $a$ . However, there is one important difference, in as much as the integral over  $f(\epsilon)$  has a singularity caused by  $\alpha = 1.5$  and 2.5 (corresponding to a Boltzmann parameter  $n = -0.5$  and  $n = -1.5$ ) as  $\epsilon$  approaches zero. The integral between a lower limit of  $\epsilon$  and infinity is sensitive to the chosen value of the lower limit and therefore is not too useful, at least for an integral representation as needed in the Lorenz curve.