Article

Elementary Principles in Statistical Economics

Cal Abel [1,*]

[1] Signal Power and Light, Inc.; crabel@signalpowerandlight.com
* Correspondence: crabel@signalpowerandlight.com

Abstract: Economics has long sought to bridge the principles of microeconomics into the realm of macroeconomics. This paper presents a formal attempt to do so by using a maximum entropy based approach derived from statistical mechanics coupled with subjective game theory and elements from political philosophy. This approach is then applied to income distributions and to the Cobb-Douglas production function, to create a framework for future applications, and to illustrate where past work had made implicit assumptions regarding the system. The paper then explores the consequences of the approach, illustrating self-contradictions in the political philosophy of distributive justice, formally deriving an equation of state for the transactions on the Bitcoin network, and deriving from this the ideal gas law and polytropic process for an economy proving that expansionary monetary policy is extractive – not stimulative.

Keywords: statistical mechanics; information theory; game theory; subjective utility; entropy; income inequality; distributive justice; monetary policy;

---

1. Introduction

Approximately fifty-years before Claude Shannon developed information theory, Josiah Willard Gibbs used an equivalent mathematical approach to formally derive equilibrium thermodynamics.[1] In his seminal work, Gibbs showed that the expected uncertainty of information theory is mathematically equivalent to Clausius' entropy. Gibbs began with Hamilton's equations of motion to describe each individual particle and then derived the equation of state using an information theoretic approach [1]. The power of Gibbs' approach allowed the seamless integration of quantum mechanics, while Boltzmann's approach had trouble incorporating quantum mechanics. Using Jaynes' simplification of Gibbs' method [2], this paper will show how one can from the micro-principles of game theory formally derive a macroeconomic model, an equation of state of the system of individual economic actors.

In economics, it is generally considered to be impossible to aggregate from microeconomics into macroeconomics. Many critiques of macroeconomics center around the loss of the individual action in the group due to the reductiveness of the aggregation [3] (Ch. 12). This paper shares the same critique and shows that the reductiveness of a common classical model is fundamentally the neglect of an entropic term that arises from the allowed manifold of individual action and choice. Essentially, if a macroeconomic model does not explicitly state an entropic term, that model is making an implicit assumption about the evolution of entropy of the system that ignores the action of the individual actors within that system.

In classical thermodynamics, the fundamental unit of account is the atom. Where classical thermodynamics diverges from quantum mechanics is primarily in the subatomic scale. The classical approach provides a more accessible and usable model with limited loss of accuracy for most practical aspects of statistical mechanics, e.g. modeling the spatial and temporal neutron flux of a nuclear reactor. Using a similar line of argumentation, this paper takes the approach to derive macroeconomics relying on the individual as the fundamental unit of account using classical (continuous) distributions. It is not technically correct but is a "good enough" approximation.

There is a relatively recent field of econophysics which uses the mathematics of thermodynamics to describe economical systems. To the author's knowledge, a formal justification of this approach has not been offered. This paper attempts to rectify this problem and to formally establish the field of statistical economics.

2. Macroeconomic Equation of State

2.1 Game Theory and Subjective Utility

To begin the construction of statistical economics, we begin with John von Neumann and Oskar Morgenstern's axiomatic derivation of game theory [4]. Their expression of the outcomes of the game is as being inherent properties of the game and the players. Their ergodic approach was relaxed in 1968 by J. Pfanzagl to allow for subjective measurement of the outcome of games [5] (§12.6). "The essential relaxations are allowing for divergence between objective and subjective probability, and not requiring the range of probabilities to be connected" [5] (p. 219).

Ellsberg's [6] and Allias' [7], critique of Savage's subjective expected utility theory is correct, as Savage's maximization of utility results in paradox and contradiction when individuals made choices. Taking Ellsberg's game as an example [5] (pp. 653-654), Table 1. Summary of Ellsberg's Game, we can see two different preferences for risk. Ellsberg noted that people in a greater fraction would prefer the set of the shaded wagers (I,IV) over the set of unshaded wagers (II,III). But that the pairings were always the same.[6] Ellsberg [6] (p. 655) and Pfanzagl [5] (p. 206) saw this as a contradiction to the Savage's lack of illusion principle. In their minds, the ordering should be (I,III) and (II,IV) as the outcomes under uncertainty represent virtually identical wagers. Assuming an equal probability of black or yellow, this would be true, but the issue is not about the expectation, it is about the entropy of the outcome.

Table 1. Summary of Ellsberg's Game [5] (pp. 653-654) where $\alpha \in [0,60]$

	Odds	30	α	$60-\alpha$			
		Payout					
Game	Wager	red	black	yellow	Expectation	Range	Entropy
A	I	100	0	0	$100 \cdot \frac{1}{3}$	$100 \cdot \frac{1}{3}$	0.366
	II	0	100	0	$100 \cdot \frac{\alpha}{90}$	$0 \to 100 \cdot \frac{1}{3}$	0.304
B	III	100	0	100	$200 \cdot \left(1-\frac{\alpha}{90}\right)$	$200 \cdot \frac{1}{3} \to 200$	0.245
	IV	0	100	100	$200 \cdot \frac{2}{3}$	$200 \cdot \frac{2}{3}$	0.270

The participants in the game consistently selected the outcome which resulted in the greatest entropy. It is not just utility maximization that they are doing. It is also a function of the entropy maximization.[8] This does not explain why there are still a set of people who choose (I,III), here we can turn to the entropies for an explanation. We can estimate $P(I,IV) = e^{-(s_I+s_{IV})+(s_{II}+s_{III})} = 91.6\%$.

Because "[h]uman action is necessarily always rational" [3] (p. 19), we need to rethink defining rational behavior as the "sure-thing principle and the lack of illusion principle" [5] (p. 205). Interestingly, the example cited by [5] (p. 205) as an example of how measured outcomes don't match rational behavior, is the example illustrated in Table 1, where it was shown that Ellsberg and Pfanzagl were maximizing the wrong thing. For this reason and the repeated hang-up that occurs on defining utility, it will be necessary to abandon subjective utility to focus on subjective probability.

2.2 Subjective Probability

In order to allow for each individual to act without passing judgement on their actions as to what is and is not rational, we need to provide a minimal definition for quantifying their action and to root the definition of rationality broadly. First is to adopt von Mises' requirement that all action be rational [3] (p. 19). To do this we state that all action that is theoretically observable is rational action – that we restrict ourselves to define rational action as demonstrated action. Second, to quantify the observable action, we rely upon Savage's quantitative personal [subjective] probability [9] (pp. 33-40) to express our knowledge of that action.

2.3 Maximum Entropy Estimates

By adopting Pfanzagl's axioms [5] (p. 213), we can express our knowledge of the outcome from a finite set of options, with no knowledge of the actual underlying utility function. At this point our knowledge of the individual's choices can be expressed as Savage's personal probability density function $f(x)$ [5] (p. 220). Following Jaynes [2] (p. 623), define the probability density function as:

$$f(x) = e^{-\sum_i \beta_i \varphi_i(x)}, \tag{1}$$

where, $\varphi_i(x)$ are the moments of $f(x)$, $\varphi_0(x) = 1$, and where β_i are the Lagrangian multipliers that maximize the information entropy of the distribution with the system subject to the constraints,

$$\mathbb{E}[\varphi_i\,(x)] = \int dx\, \varphi_i\,(x) f(x) = a_i. \tag{2}$$

The expectation of the 0th moment must be 1, $a_0 = 1$. This is Jaynes' normalization constraint.[2] It is a more general statement of Gibbs' *conservation of the extension-in-phase* [1] (p. 10). The normalization constraint directly results in the partition function,

$$Z(\beta) = e^{\beta_0} = \int dx\, e^{-\sum_{i>0} \beta_i \varphi_i(x)}. \tag{3}$$

The remaining Lagrangian multipliers can be determined from,

$$\mathbb{E}[\varphi_i\,(x)] = -\frac{\partial}{\partial \beta_i}\ln(Z) = -\frac{\partial \beta_0}{\partial \beta_i}. \tag{4}$$

The resulting maximum entropy estimate of the individual's demonstrated utility is,

$$s = \sum\nolimits_i \beta_i a_i. \tag{5}$$

So far, there has been no loss in generality in deriving equation (5). There are other ways of arriving at equation (5) that do not rely on using moments, notably reference [10].

2.4 Group Entropy

Before proceeding with the group aggregation, we need to assume logical independence of each individual's utility. This assumption has a physical meaning of the freedom of individual choice – there is no mind control. It is another way of stating Rawls first principle of justice that, "[e]ach person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberties for all"[10] (Ch. V Sect. 46). This assumption is a maximum entropy assumption as any forced or coercive acts will reduce the choice of action of the individual, thus the allowed states and associated individual complexity.[3] As Jaynes notes, the assumption of the principle of maximum entropy is Laplace's "principle of insufficient reason" [2] (p. 622).

Mathematically this results in the group density function, $F(X)$, being separable in $x_j\ \forall\ x_j \in X$.

$$F(X) = \prod_j f_j(x_j). \tag{6}$$

By definition,

$$S = -\int dX\, F(X) \ln[F(X)]. \tag{7}$$

Due to the normalization constraint and the logical independence of the actors, equation (7) reduces to,

$$S = \sum_j s_j. \quad (8)$$

At this point, no assumptions have been made outside of the nature of the individual utility other than each individual's expression of freedom to choose. Thus, equation (8) formally represents a *complete* group utility function. It is not however very tractable. To make it usable we need to assume Rawls' veil of ignorance. "This ensures that no one is advantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances" [10] (Ch.1 Sect. 3). The veil of ignorance needs to be extended as Rawls did not take the it far enough. The veil needs to be extended across time. What this means is that not only does one not know who they can choose to be in a society, but that they also cannot pick at what point in an individual's life they can choose. This interpretation of the veil effectively states that each of us given a similar set of circumstances and experience would make a similar set of choices and results in,

$$S = Ns, \quad (9)$$

where, N is the group's population.

The assumption of maximum entropy as a constraint provides some useful insights into the nature of a society. First is that social rules are constrained – rules cannot arbitrarily provide an individual access to additional choice without providing the necessary means exogenously. Thus, endogenous rules cannot create out of thin air, they can only restrict action. These social rules that have evolved over time affect the social entropy. For example, if there are any systemic effects where one group is disadvantaged to another, what will show up is that there will be a reduction in the total social entropy from those rules which provide maximum entropy subject to the endogenous and exogenous constraints: resources, history, technology, warfare, social reproductive needs, etc.

What this provides the political scientist and economist is the ability to quantify the impact of a set of different policies. Those policies which increase social entropy should be kept and those which do not, discarded. This is a consequence of the second law of thermodynamics. Social change will occur spontaneously if the entropy gradient is positive. If the entropy gradient is negative, the society will have to consume additional resources to enforce the rule. One could say that we act as if "…led by [entropy] to promote an end which was no part of [our] intention" [12] (Book IV Ch. II).

There are cases such as national defense, when a reduction in social entropy occurs in order to fund the defense of a nation. This slight reduction helps to provide protection against exogenous "black swan" events, which can act to significantly reduce social entropy. But here the appropriate amount of reduction in individual liberty to promote general liberty is not known and has to be determined at the social level. The justification of reducing liberty to promote liberty is Rawls' first priority rule [10] (Ch. V Sect. 46) and is warranted.

3. Quantitative Applications

There are two practical examples that deserve exposition. The first is the application of the theory applied to income distributions and the second is to derive a Cobb-Douglas like production function from the multivariate log-normal distribution.

3.1. Income Distributions

A society is group of individuals. Corporations employ individuals and use capital to produce goods and are owned entirely by individuals. Using this reasoning, by exploring individual income we are seeing the net impact of all economic action within a given society. This has long been understood in economics as being important, e.g. Vilfredo Pareto whom the Pareto distribution was named after for his work using it to study income distributions.

In a more modern physics-based Banerjee and Yakovenko modeled United States Adjusted Gross Income, AGI, using a stationary Fokker-Plank equation to combine an exponential and Pareto distributions [13] (p. 10). The exponential distribution modeled the portion of the economy in thermal equilibrium, where the interactions are additive [13] (p. 10). The Pareto distribution modeled the epithermal region where interactions are multiplicative and are associated with returns on capital [13] (pp. 10-11).

3.1.1. Physical Analogies

The new proposed model has a direct physical analogy with neutron thermalization. If we consider the idea of financial interactions with each other as predominantly being an equivalent exchange of value for goods/services, we have a direct analogy with physics through elastic collisions such as colliding billiard balls or neutrons interacting with atoms. The statistics of neutron interactions are significantly more worked out than billiard balls, so we will turn our attention to neutron thermalization.

Neutrons are born with an average energy of ~ 2MeV. Most of the reactors in the world rely on neutrons that have slowed down to thermal energies and generally use water (hydrogen) as the primary means of slowing down the neutrons. This results in three regions of neutrons: the source region, the slowing down region, and the thermal region. The neutron energy profile $\propto 1/\mathrm{E}$ in the slowing down region [13,14].

In the thermal region, the neutrons reach thermal equilibrium with the moderator with an average energy of ~0.025 eV [15] (p. 379). The neutrons adopt a Maxwellian distribution in this region [13,14].

The mechanism that drives the thermal and epithermal distributions is elastic scattering, where energy and momentum are conserved in the laboratory frame. The epithermal region is generally characterized by trivial up-scatter (gaining energy) probabilities and is entirely dominated by down-scatter (loosing energy). The thermal region is when there is a non-trivial probability for the neutron to gain energy from scattering. Mathematically for particle interactions of the same mass this can be expressed with the following transfer probability for up-scatter, $E' \leq E$, and down-scatter, $E' > E$:

$$\sigma_s(E')f_s(E' \longrightarrow E) = \begin{cases} \frac{\sigma_{s0}}{E'}\exp\left(\frac{E'-E}{kT}\right)\mathrm{erf}\left(\sqrt{\frac{E'}{kT}}\right) & for\ E' \leq E \\ \frac{\sigma_{s0}}{E'}\mathrm{erf}\left(\sqrt{\frac{E}{kT}}\right) & for\ E' > E \end{cases} \quad \text{[13] (p. 337).} \tag{10}$$

The up scattering is statistically significant up to about energies of $E' \leq 25kT$ [14] (p. 336).

Assuming no loss through absorption we have the following integral equation:

$$\sigma_s(E)p(E) = \int_0^\infty dE'\sigma_s(E')f_s(E' \longrightarrow E)p(E') + S(E) \text{ [14] (p. 378).} \tag{11}$$

There are two limiting solutions to the above integral equation that provide a closed form first is no up-scatter well above equilibrium and the second is for neutrons in thermal equilibrium with no source term. The two solutions result in the following limiting scenarios respectively:

$$p(E) \propto \frac{1}{E}, \tag{12}$$

$$p(E) = E\beta^2 e^{-E\beta}, \tag{13}$$

where, $\beta = {}^1/_{kT}$ [15] (p. 378). Following Gibbs' "On Certain Important Functions of the Energies of a System", his equation (290) provides the distribution of a particle's energy as $\sim \Gamma({}^n/_2, \beta)$, where n is the number of degrees of freedom of the system [1] (p. 93). Thus, the Maxwell distribution represents a system of 4-degrees of freedom and the Gibbs distribution used by Banerjee and Yakovenko [13] represents a system of 2-degrees of freedom.

3.1.2. Economic Data Model

It is useful to continue the analogy of an individual's income as being analogous to a particle's energy. Banerjee and Yakovenko found that the thermal population of Adjusted Gross Income in the United States followed,

$$p(E) = \beta e^{-E\beta} \text{ [13] (p. 10).} \tag{14}$$

and that the combined Pareto/exponential distribution was

$$P(r) \propto \frac{e^{-r_0\beta \arctan\left(\frac{r}{r_0}\right)}}{\left(1+\left(\frac{r}{r_0}\right)^2\right)^{\frac{1+\alpha}{2}}} \text{ [13] (p. 11).} \tag{15}$$

They derived equations (14) and (15) using stationary solutions to the Fokker-Plank equation,

$$\frac{\partial P(r,t)}{\partial t} = \frac{\partial}{\partial r}[A(r)P(r,t)] + \frac{\partial^2}{\partial r^2}[B(r)P(r,t)]. \tag{16}$$

Following their methodology, the coefficients,

$$A(r) = \frac{A_0}{r}, \qquad B(r) = \frac{B_0}{r}, \tag{17}$$

will result in a Maxwell distribution for the stationary solution to equation (16)

$$P(r) = r\beta^2 e^{-r\beta}, \qquad \frac{1}{T} = \beta, \qquad \beta \equiv \frac{A_0}{B_0}. \tag{18}$$

Recalling from reference [13] (p. 10) the epithermal asymptote:

$$A(r) = a\,r, \qquad B(r) = b\,r^2, \qquad \alpha \equiv 1 + \frac{a}{b} \tag{19}$$

and combining it with equation (17) will result in the combined drift and diffusion coefficients of,

$$A(r) = \frac{A_0}{r} + a\,r, \qquad B(r) = \frac{B_0}{r} + b\,r^2, \qquad {r_0}^3 \equiv \frac{B_0}{b}. \tag{20}$$

This will result in a stationary solution to equation (16) of:

$$P(r) \propto \left(\frac{1 - \frac{r}{r_0} + \left(\frac{r}{r_0}\right)^2}{\left(1 + \frac{r}{r_0}\right)^2}\right)^{\frac{r_0\beta}{6}} \frac{r}{\left(1+\left(\frac{r}{r_0}\right)^3\right)^{\frac{2+\alpha}{3}}} e^{-\frac{2}{\sqrt{3}} r_0\beta \arctan\left(\frac{1-\frac{2r}{r_0}}{\sqrt{3}}\right)}. \tag{21}$$

To estimate the parameters a hierarchical Hamiltonian Markov Chain Monte Carlo with No U-Turn Sampling, HMC with NUTS. The parameter prior distributions were assumed to be $\sim \Gamma(\alpha_\Gamma, \beta_\Gamma)$ and are provided in Table 2. σ is the standard deviation of the log-normal distribution of the error between the data and the model. The Stan code, all of the source data, and the Rstudio™ project implementing the Stan code can be found at https://github.com/crabel99/incomeDist.

Table 2. Prior distribution hyperparameters

Parameter	α_Γ	β_Γ
σ	2	$8{\cdot}10^{-6}$
T	4	$8{\cdot}10^{-5}$
r_0	4	$4{\cdot}10^{-5}$
α	2	1

3.1.3. Economic Modeling Results

The data was taken from IRS Publication 1304 Tables 1.1 and 2.1 for all available tax years, 1996-2017 and 1993-2017 respectively [16]. Because of the relatively small sample size, ~20 bins/tax year, the compensated Akaike Information Criteria, AICc, was used to evaluate the goodness of fit of the different models,

$$AICc = \frac{2k\,n}{n-k-1} - 2\ln(\hat{L}), \tag{22}$$

where n is the number of datapoints, k is the number of coefficients and $\hat{L}$ is the maximum of the likelihood function [17].

The use of AGI income for modeling income distributions, does not lend itself to making a clear understanding of the data. In the publicly available dataset, the only table which included anything related to Gross Total Income, GTI, was IRS Table 2.1 [16], which was for the returns with itemized deductions. In that data set, the AGI, Total Income, and the Taxable Income roughly follow each other. IRS Table 1.1, which was used by Banerjee and Yakovenko [13], only has AGI and taxable income. Because of the rough correlation between Total Income and Taxable Income in IRS Table 2.1, the taxable income was used as a proxy for Total Income.

Table 3 provides the AICc for comparing the modified Gibbs distribution, equation (15), and the modified Maxwell distribution, equation (21), for each of the three datasets: AGI from IRS Table 1.1, Total Income from IRS Table 1.1, and AGI from Table 2.1. For the AGI in Table 1.1 equation (15) is clearly the best fit. For the Total Income in Table 1.1 and the AGI in Table 2.1, equation (21) provides the best fit.

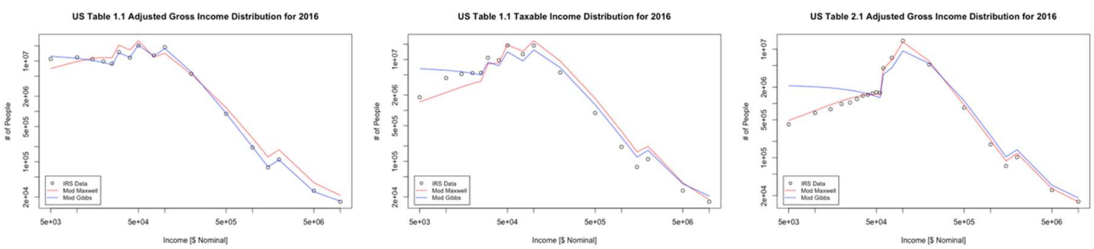
Table 3. AICc for the different models and datasets. When comparing models within a dataset, the lower AIC is the better fit.

	Table 1.1 AGI		Table 1.1 Total Income		Table 2.1 AGI	
Year	Gibbs	Maxwell	Gibbs	Maxwell	Gibbs	Maxwell
1993					55.4	13.4
1994					54.7	9.8
1995					53.6	6.1
1996	-42.0	57.7	53.4	19.7	55.1	5.8
1997	-41.1	56.2	51.0	19.8	55.4	7.7
1998	-36.5	56.3	66.3	11.0	56.3	4.9
1999	-45.2	57.6	64.4	8.6	56.4	2.0
2000	-51.8	61.2	72.1	-5.6	62.0	8.5
2001	-48.4	64.3	84.5	-1.4	59.0	2.1
2002	-39.8	65.0	94.5	15.0	57.1	0.4
2003	-20.9	21.5	42.5	-8.5	53.5	-11.0
2004	-21.7	22.4	43.1	-6.9	54.3	-12.7
2005	-25.1	22.1	41.4	-12.1	53.8	-14.4
2006	-27.8	23.8	41.1	-17.2	54.1	-13.1
2007	-29.6	24.1	40.3	-22.8	54.4	-13.7
2008	-34.5	24.8	47.0	-1.7	50.7	-13.2
2009	-32.7	24.2	52.6	16.0	48.7	-14.0
2010	-30.6	23.5	53.1	17.8	51.5	-13.6
2011	-34.8	24.8	52.6	19.0	51.2	-13.4
2012	-31.8	23.4	55.2	22.4	51.7	-17.6

2013	-38.3	25.7	55.2	20.7	52.8	-8.6
2014	-38.6	25.3	55.3	21.1	54.2	-6.9
2015	-40.5	25.5	57.4	25.3	55.7	-9.0
2016	-40.6	25.6	59.0	28.1	54.8	-9.6
2017	-41.1	25.5	58.2	25.5	56.3	-16.2
Ave	**-36.1**	35.5	56.4	**8.8**	54.5	**-5.1**

The above result suggests that AGI from Table 1.1 is a problematic measure as the cutoff of what is a standard deduction is entirely arbitrary. It is not measuring the value that society places on the action of an individual (whether that is deserved or not is not what is being asked). From this perspective GTI would provide a more accurate measure.

Figure 1 provides a visual representation of the data from 2016, which based on the AICc, provides the least amount of evidence to support the methodology of using Taxable Income as a proxy for GTI.



(a) (b) (c)

Figure 1. Sample plots of the modified Gibbs distribution, equation (15), and the modified Maxwell distribution, equation (21), fit to the data from **(a)** the AGI from IRS Table 1.1 [16], **(b)** the Taxable Income from IRS Table 1.1 [16], and **(c)** the AGI from IRS Table 2.1 for the 2016 tax year [16].

The reason for the focus in the modeling on the GTI is that it is the sum of all money transferred to an individual in a year. For the purpose of assessing the value of an individual's contribution to society under the veil of ignorance, GTI would represent this utility the closest. If we wanted to look at the impact of redistributive policies, GTI would need to be modified by removing taxes paid and adding transfers. Using the methodology derived in Section 2 as a normative basis, the two policies, the null and the redistributive, could be quantitively compared. Because the author did not have access to the needed underlying data, such a comparison could not be made at this time. The purpose of this modeling was to conceptually test the methodology and the distribution needed for assessing the income data.

3.2. Generalized Cobb-Douglas Production Function

The Cobb-Douglas production function is a common production function, equation of state, used in macroeconomics. Its general form is

$$Y(t) = A(t)\, K(t)^{\alpha} L(t)^{1-\alpha}, \qquad \alpha \in [0,1]. \tag{23}$$

It is treated as an input-output model where the inputs K, capital, and L, labor, have a constant return to scale of Y, typically Gross Domestic Product, GDP [18] (p. 184).

This model has had some adaptations where $A(t)$, is not constant with time and is commonly referred to as the Solow residual or Total Factor Productivity and is generally attributed to "technological progress". Ayers and Warr showed that this residual was not technological progress, but was instead the exergetic, useful work, input into the economy. The proposed a modification to equation (23), which acknowledged the first law constraints all action is subjected to,

$$Y(t) = A(t)\, K(t)^{\alpha} L(t)^{\beta} E(t)^{\gamma}, \qquad \alpha + \beta + \gamma = 1, \tag{24}$$

where E is the payment to exergy [18] (p. 185).

They showed that the Solow residual was primarily accounted for by exergy. From a physical sense, this is fully justified. Take the analogy of a person digging a hole. If they dig with their hands, there is not input from capital, they are the labor, and the food they eat is converted into exergy. If given a shovel, they now have the capital input of the shovel, they remain the source of labor, and their food is the sole input for exergy. Now, provide them with an excavator. The capital input is the lease paid to rent the excavator, the wage that they are paid is their labor, and the exergy input comes from the diesel fuel consumed by the excavator.

In each of these three examples the amount of work done in a given time increases with a constant input of labor. In the first two examples the exergy input is the same, while the capital input changed, making the work more efficient from a labor and exergetic perspectives. In the third example the capital input is drastically increased, but what drives the capital controlled by the labor is the combustion of the diesel fuel in the excavator. It is not an endogenous factor as what Solow originally attributed it to be. It is simply a first law balance. Cobb, Douglas, Solow, and many other economists ignored that the economic world is embedded in the physical world and therefor is bound by the same underlying constraints. Ayers and Warr pointed out the lack of a simple acknowledgement of the first law constraint.

They did acknowledge the second law in their book; however, they did not fully incorporate the second law in their modeling. They did account for the second law in adjusting the economic input of energy to be exergy, useful work.[18] (pp.137-140) This was as far as they took the second law. They did not explicitly look at the distributional aspect of entropy.

3.2.1. Expanding the Univariate Log-normal Distribution

We begin by setting the moment vector for the univariate case.

$$\varphi(u) = \begin{bmatrix} 1 \\ \ln(u) \\ \left(\ln(u)\right)^2 \end{bmatrix}. \tag{25}$$

Placing equation (25) in equation (3) results in the log of the partition function as

$$\beta_0 = \frac{1}{2}\ln\left(\frac{\pi}{\beta_2}\right) + \frac{(\beta_1 - 1)^2}{4\beta_2}. \tag{26}$$

With a corresponding gradient of

$$-\nabla\beta_0 = \begin{bmatrix} 1 \\ \dfrac{(1-\beta_1)}{2\beta_2} \\ \dfrac{1}{2\beta_2} + \dfrac{(\beta_1 - 1)^2}{4{\beta_2}^2} \end{bmatrix}. \tag{27}$$

We set the moment constraint as

$$\mathbb{E}[\varphi(u)] = \begin{bmatrix} 1 \\ \mu \\ \mu^2 + \sigma^2 \end{bmatrix}, \tag{28}$$

and solving equation (4) results in the Lagrangian multiplier vector of

$$\beta = \begin{bmatrix} \dfrac{\mu^2}{2\sigma^2} + \dfrac{1}{2}\ln(2\pi\sigma^2) \\ 1 - \dfrac{\mu}{\sigma^2} \\ \dfrac{1}{2\sigma^2} \end{bmatrix}. \tag{29}$$

Recalling,

$$\langle u\rangle = e^{\mu+\frac{\sigma^2}{2}}, \tag{30}$$

the entropy becomes,

$$\mathrm{s} = \frac{1}{2}\ln\left(2\pi e\sigma^2 e^{-\sigma^2}\right) + \ln(\langle u\rangle), \tag{31}$$

Looking at the extrema of equation (31) shows that the entropy is unbounded for $\langle u\rangle$ and unbounded negative for σ^2 as it approaches 0 or ∞. Note, the boundary condition of Rawls theory of distribution justice, which "…holds that social and economic inequalities, for example inequalities of wealth and authority, are just only if they result in compensating benefits for everyone, and in particular for the least advantaged of society" [10] (Ch. 1 Sect. 3) represents the case of $\sigma^2 \to \infty$. There is however a maximum entropy for $\sigma^2 = 1$,

$$\frac{\partial s}{\partial \sigma^2} = e^{-\sigma^2}(1-\sigma^2) = 0, \tag{32}$$

which corresponds to a $Gini \approx 0.531$. For the log-normal distribution, σ^2 is the only parameter that affects the distribution's shape.

3.2.2. Variable Transformation

We can set a new variable, ϵ, where $u = \alpha\ln(\epsilon)$. This variable is lognormally distributed and has an expectation of $\langle\epsilon\rangle = e^{\frac{\mu}{\alpha}+\frac{1}{2}\left(\frac{\sigma}{\alpha}\right)^2}$, where μ and σ are defined from before the transformation. Substituting the new transformed expectation into equation (31), results in,

$$\mathrm{s} = \frac{1}{2}\ln\left(2\pi e\sigma^2 e^{-\left(\frac{\sigma}{\alpha}\right)^2}\right) + \alpha\ln(\langle\epsilon\rangle). \tag{33}$$

3.2.3. Multivariate Log-normal Distribution

We continue with the methodology of the univariate case and apply it to the multivariate case:

$$\boldsymbol{X} = [x_1 \quad \cdots \quad x_k], \tag{34}$$

$$\varphi(\boldsymbol{X}) = \begin{bmatrix} 1 \\ \ln(\boldsymbol{X}) \\ \ln(\boldsymbol{X}^T)\ln(\boldsymbol{X}) \end{bmatrix}, \tag{35}$$

$$\mathbb{E}[\varphi(\boldsymbol{X})] = \begin{bmatrix} 1 \\ \boldsymbol{\mu} \\ \boldsymbol{\mu}^T\boldsymbol{\mu} + \boldsymbol{\sigma}^T\boldsymbol{\sigma} \end{bmatrix}, \boldsymbol{\mu} = [\mu_1 \quad \cdots \quad \mu_k], \boldsymbol{\sigma} = [\sigma_1 \quad \cdots \quad \sigma_k], \tag{36}$$

which results in an entropy of

$$\mathrm{s} = \frac{1}{2}\ln\left((2\pi e)^k|\boldsymbol{\Sigma}|e^{-Tr(\boldsymbol{\Sigma})}\right) + \sum_{i=1}^{k}\ln(\langle x_i\rangle). \tag{37}$$

A similar transformation can be done from Section 3.2.2 with equation (37)

$$\mathrm{s} = \frac{1}{2}\ln\left((2\pi e)^k|\boldsymbol{\Sigma}|e^{-Tr(\boldsymbol{\Sigma}\oslash\mathbf{A})}\right) + \sum_{i=1}^{k}\alpha_i\ln(\langle x_i\rangle), \mathbf{A} = \boldsymbol{\alpha}^T\boldsymbol{\alpha}, \boldsymbol{\alpha} = [\alpha_1 \quad \cdots \quad \alpha_k]. \tag{38}$$

where $\oslash$ is the elementwise Hadamard divisor. Equation (38) has the same mathematical form as the equation of state of an ideal gas [19], it is also mathematically equivalent to (24) under the isentropic condition.

As engineers, chemists and physicists are well familiar with, isentropic transformations represent an idealized situation. It is an assumption that does not hold in the observed world. Economists in general have made the assumption of ignoring that they are working with distributions of outcomes, where in reality each decision is perfectly rational and plausible, not just the average.

When we formally take into account the distribution of outcomes resulting from individual choice, the entropy term appears and we see the implicit assumption of the economists.

3.3. Selecting a More Absolute Deflator

To make comparisons across years, some measure needs to be taken into account to account for the fluctuations in the value of a currency. This can be done by selecting a "basket of goods" such as the Consumer Price Index or by referencing the value of the currency to the GDP of the main country such as the GDP deflator. In the author's mind these are entirely arbitrary definitions that don't sufficiently tie the value of the currency to action. In the physical world the currency that is exchanged is energy. If exergy is roughly 80% of GDP [18], then why not measure the value of the currency relative to energy, setting the marginal utility of money in terms of GJ/$?

3.3.1. Utility in Bitcoin[1]

Bitcoin represents an interesting use case as it is a functioning global currency. As of September 21, 2020, it is the 6th largest currency globally by market capitalization [20]. When the author originally became involved in Bitcoin, it was still an experiment of "collectible tokens". It can no longer be considered an experiment, and closer study will help elucidate some facts about currencies in general and further support grounding the value of a currency in energy.

Bitcoin, at its very base level, is simply a database that is distributed globally, with very specific rules for how updates in the ledger are made. The database is updated in roughly 10-minute intervals through a process known as mining [20]. These roughly 10-minute intervals, known as blocks, are the accepted and validated transactions that were broadcast to the network in that period of time [20]. A "block" is a block of database entry updates. Thus to reconstruct the database, you have to start from the genesis block and track every transaction to date. This is just a set of all blocks, it is not yet a "blockchain".

To secure the network and to prevent against the Byzantine General problem, Satoshi Nakamoto incorporated the idea of chaining the hashes and providing a mechanism for rewarding the miner who presents the network with the greatest "proof of work" [20]. The process of chaining the hashes is what creates the blockchain. Here is how mining works:

1. The miner will select the transactions from a validated pool of transactions that it wishes to include in the next block [20].
2. It then structures these transactions into a data header, along with the hash of the previous block, and a random number that the miner generates called a nonce [20].[2]
3. This data header is then hashed using a SHA-256 algorithm. For the hash to be accepted by the network it needs to have sufficient difficulty, enough preceding zeros in the hash to conform with the difficulty rules. If not, it is rejected [20].

What the miners fundamentally do is to guess random numbers and then generate hashes of the random numbers they just generated with the data to be included in the block, and the header of the previous block. Currently, this is done ~ 130 quintillion times a second globally [22]. The difficulty parameter is what is adjusted to control the block issuance time to keep it at roughly 10 minutes [20]. Using the difficulty and the time to find a new block one can calculate an estimate of the global hash rate for the network.

This hash rate is the "on-chain" measure of what the value of bitcoin is, as that amount of computation is being expended to stochastically compete for the block reward. The on-chain value of bitcoin is measured in terms of information in an arbitrary frame of reference. To root it into the

[1] The descriptions in this section come predominantly from [21] unless otherwise noted.

[2] This is not all of the data that is included but is enough to provide the needed explanation here. For those who are more interested in this process see [21] or look at the source code: https://github.com/bitcoin/bitcoin

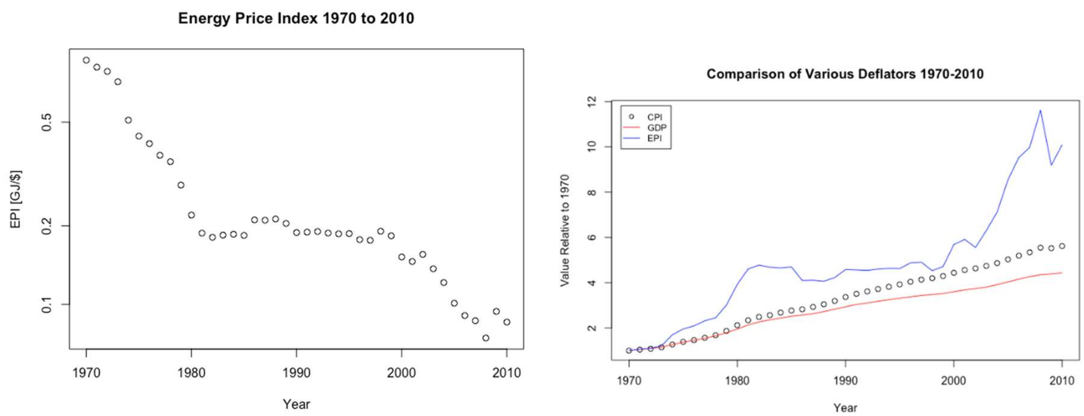
physical world we can make an estimate of the average energy expended per hash, Joule per tera hash, J/TH.

If we were to assume only the most efficient miners commercially available, ~30 J/TH, the resulting global power consumption would be at a minimum 3.9 GW at the current difficulty, resulting in a minimum marginal utility of bitcoin at ~0.37 GJ/mBTC or 102 kW-hr/mBTC. With current market prices, $10.79/mBTC, this is equivalent to a minimum marginal utility of the dollar of ~0.035 GJ/$. This corresponds to a theoretical maximum network value of power of $105/MW-hr. The miners then make the difference on the spread of the cost of electricity to mine and the miner's maximum value of power. Fundamentally, miners are converting energy into a secure database.

Because this is done programmatically and in an entirely trustless manner using cryptographic hash functions, Bitcoin represents a compact and concise summation of the different elements of what makes a currency. In measuring fiat currencies, they have an entirely different set of rules and their value is not set programmatically. However, we are not unjustified in thinking of valuing a currency in terms of energy, as the 6th largest currency in the world has been since its creation.

3.3.2. Energy Price Index

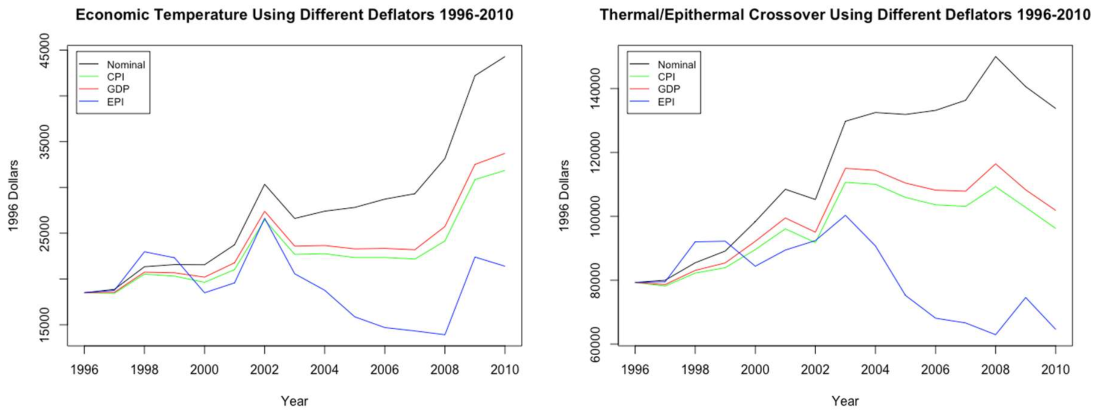
In order to develop the measure of energy for a currency, the author looked at data from the Energy Information Administration, EIA, which tracks the amount of energy sold each year from each fuel source along with how much was spent on that energy [23] (Table 1.5). This is from data aggregated from utilities, refineries, and distributors across the nation [23]. Figure 2 shows the estimated Energy Price Index, EPI, from EIA data [23] (Table 1.5), the GDP deflator [24] and the Consumer Price Index- Urban Consumers CPI-U [25].



(a) **(b)**

Figure 2. (a) The Energy Price Index derived from [23] (Table 1.5) plotted on a logarithmic vertical axis. This represents a measure of the marginal utility of the dollar. **(b)** The Consumer Price Index [25], Gross Domestic Product Deflator [24], and the inverted EPI from (a). All indices are normalized relative to their 1970 values.

Using the different deflators from Figure 2(b) to adjust the economic temperature and the thermal/epithermal crossover from the earlier regressions of the US income distribution using the Taxable Income from [16] (Table 1.1) can be seen in Figure 3. Banerjee and Yakovenko [13] (p. 13) noted that during periods of monetary expansion the fraction of the society in the epithermal region tended to increase. This is true only in nominal terms. In real terms, the impact of monetary expansion removes the apparent gains.



(a) **(b)**

Figure 3. (a) The estimated economic temperatures derived from the Maxwell model of the [16] (Table 1.1) Taxable Income data adjusted using the different deflators from Figure 2. The deflated temperatures were normalized to the nominal temperature in 1996. (b) The estimated thermal/epithermal crossover, r_0, derived from the Maxwell model of the [16] (Table 1.1) Taxable Income data adjusted using the different deflators from Figure 2. The deflated crossovers were normalized to the nominal crossover in 1996.

The adjusted parameters from Figure 3 were numerically integrated to provide the specific social entropy, Figure 4. In interpreting the results, an increasing social entropy is stating that the individual has more allowed states to be able to choose from – an increase in social choice/complexity. Conversely, a reduction in social entropy is a reduction in the social complexity of the individual. If we adopt the EPI as a deflator, we see the 2000 recession from the dot com bubble, with a "V" shaped recovery in 2002. After 2002, the US entered a protracted recession culminating in 2008. There was a noticeable recovery in 2009. Furthermore, when using the EPI, the colloquial statements of the economy "heating up" or "cooling down" are technically accurate.

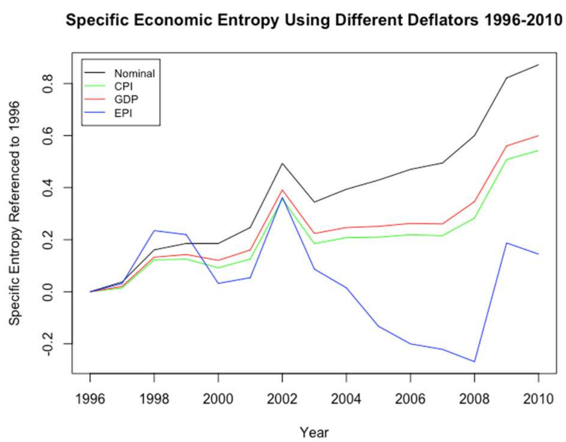


Figure 4. The specific social entropy using the Maxwell based estimation of the income distribution from the [16] (Table 1.1) Taxable Income adjusted using the different deflators of Figure 2.

4. Discussion

4.1. On "A Theory of Justice"

The reliance on subjective utility in the derivations and analysis does not preclude an objective approach, just that the mathematics will be more complex, and will reintroduce some of the paradox's solved by using the subjective approach, e.g. Alias Paradox. Because this was developed under the MAXENT subjective framework, the approach is only concerned with understanding what we observe subject to the constraint that we assume the least amount of information.

The subjective approach has some interesting philosophical implications when applied to Rawls' distributive theory of justice. Rawls used two main axioms to derive his theory of justice:

1. The veil of ignorance:
    a. "…it should be impossible to tailor the circumstances of one's own case" [10] (Ch. I Sect. 4).
    b. "…that particular inclinations and aspirations, and persons' conceptions of their good do not affect the principles adopted" [10] (Ch. I Sect. 4).
2. Justice as Fairness – "…that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles" [10] (Ch. I Sect. 4).

He sought "to look for a conception of justice that nullifies the accidents of natural endowment and the contingencies of social circumstance…" [10] (Ch. I Sect. 3) as that which would create the justice of fairness. What he saw as the outcome of wealth inequality was fundamentally an injustice:

> "Since it is not reasonable for him to expect more than an equal share in the division of social goods, and since it is not rational for him to agree to less, the sensible thing for him to do is to acknowledge as the first principle of justice one requiring an equal distribution. Indeed, this principle is so obvious that we would expect it to occur to anyone immediately.
>
> Thus, the parties start with a principle [the first principle of justice] establishing equal liberty for all, including equality of opportunity, as well as an equal distribution of income and wealth"[3] [10] (Ch. III Sect. 26).

Rawls does allow for inequality if these inequalities "… in the basic structure that work to make everyone better off in comparison with the benchmark of initial equality… The immediate gain which a greater equality might allow can be regarded as intelligently invested…" [10] (Ch. III Sect. 26). However, the ideal remains with perfect equality.

The problem is with the ideal of perfect equality, where each state has an equal probability of outcome, corresponding with a Gini coefficient of 0. What was shown earlier, is that by adopting the veil of ignorance, the inequality of outcome that occurs at the societal level is due entirely to individual choice. This distribution of choice cannot be uniform. It cannot be so because we are constrained in the natural world by the laws of thermodynamics. Thus, our set of choices is limited by the options available to us. If we were to assign equal probability to that finite set of choices, we would still be faced with the constraints of physics, which demand that we shift our attention and therefore values unequally on the world.

The idea of perfect equality can only exist in a world of infinite possibility. It is only under this condition that such an assumption is justified. Because we exist in a "post fall" world where we've gained the knowledge of good and evil, the knowledge of our own and others' limitations, we must make our choices to suit a finite reality. We have to order our individual choices, establish priorities, and create values. Not only that, that has to be done in the context of those around us and our society, and social history. It is out of this context that meaning and values emerge. While these values can be reinterpreted or deconstructed in an infinite number of ways, only a small finite set of interpretations remain valid with experience.

---

[3] Rawls is somewhat vague at this point in the book (Ch. III Sect. 26) in what he defines as his "first principle". Later (Ch. 5 Sect. 46) he separates the original first principle into the first and second principles, where the first principle becomes only the equality of liberty and the second principle becomes the equality of opportunity and the equality of outcome subject to the savings principle. He defines "[a]ll social primary goods… [as being] liberty and opportunity, income and wealth, and the bases of self-respect." [11] (Ch. V Sect. 46). He refers to these throughout the text as being where equality must focus.

If "Inequalities are permissible when they maximize, or at least all contribute to, the long-term expectations of the least fortunate group in society" [10] (Ch. III Sect. 26). The problem that we have by adopting the veil of ignorance is that the canonical inequality intrinsic in society is due to individual choice. Rawls' first principle [really only the second principle as defined later in his text] of justice is incompatible with one of the axioms he used to derive it.[4] This does not mean that everything within the first principle of justice is not supported, just the equality of opportunity and the equal distribution of income and wealth.

The equality of opportunity can be thought of as the choices that we make in life to restrict or indulge our preferences, one consequence of either restriction or indulgence is that it either opens or closes a set of space with which we can occupy. What the principle of equal liberty protects is the ability of the individuals to make those choices for themselves. In the formalism presented here, the principle of equal liberty is mathematically represented by the logical independence of each member of a society. As previously pointed out, this is also an assumption of a condition of maximum entropy. Rawls pointed out that the liberty principle must be the first principle in the "First Priority Rule", "The principles of justice are to be ranked in lexical order and therefore liberty can only be restricted for the sake of liberty" [10] (Ch. V Sect. 46). Thus, the liberty principle is in the first assumption made in the approach taken here with exploring the aggregate of the society.

The fundamental issue in Rawls' theory is that the assumption of the veil of ignorance, arrived at "naturally" very early in his text, is incompatible with the second principle of justice, "Social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality and opportunity" [10] (Ch. V Sect. 46) under the first principle – the liberty principle. These three conditions are incompatible as a set, only 2 out of the three can work together, and even then the liberty principle seems irreconcilable with the equality principle under the statistical economics framework.

The veil of ignorance is a maximum entropy assumption. It fundamentally is stating that each individual in society is indistinguishable from each other. It is only by dropping the veil that assumptions can be made about what is fair and is where prejudice is introduced. Any prejudice within society will reduce the social entropy from the ideal. The ideal distribution is the one that maximizes the entropy subject to the available constraints, which was shown to **not** be the uniform distribution. Any theory of justice must take into account that the social world is intrinsic with the physical world and subject to the same stochastic constraints of scarcity. Rawls' theory neglects this fundamental constraint. His theory and other theories that promote equality of outcome are inconsistent with the second law of thermodynamics.

On a note of the difference between the thermal and epithermal regions of income. Banerjee and Yackovenko showed that these are due to the multiplicative returns on previous savings [13] (pp. 10-11). Rawls' "just saving principle" [10] (Ch. V Sect. 44) similarly drops the veil by assuming knowledge of who in society was born where and when. A just savings principle that is consistent with the liberty principle and with the veil of ignorance is that one where policies that inherently either seek to protect or destroy the intergenerational savings should be considered unjust, because they introduce prejudice and restrict individual liberty.

Rawls rooted his theory principles of maximum entropy, the veil of ignorance, the liberty principle, and the first ordering principle. However, the equality principle can only be a valid assumption in a world with infinite possibility. The theory's seductiveness is precisely that promise of an infinite world. It is the same seductiveness of perpetual motion machines. Neither can exist in the world we inhabit.

[4] In this sentence, the first theory of justice is defined in the text above. It is not the restricted first principle that Rawls defines later as being the equality of liberty.

4.2. Production Functions and Income Distributions

One of the more surprising results was the derivation of the Cobb-Douglas production function as the isentropic version of the multivariate equation of state derived from the multivariate log-normal distribution. This result may find purchase outside of econometrics inside statistical finance, epidemiology, etc. where the log-normal distribution is commonly used.

There is evidence that the approach taken in section 2 can be integrated with the lognormal distribution [26]. Where instead of using Pareto modified Gibbs or Maxwell distributions, the income distributions would best be described as a mixture of the Pareto and log-normal distributions for the epithermal and thermal regions respectively.

Exploring the log-normal distribution some more, recall that,

$$\frac{\mathrm{Var}[u]}{\langle u\rangle^2} = \frac{\langle u\rangle^2\left(e^{\sigma^2}-1\right)}{\langle u\rangle^2} \propto \frac{1}{N} \text{ [27]}. \tag{39}$$

Add a constant of proportionality,

$$\frac{\mathrm{Var}[u]}{\langle u\rangle^2} = \frac{N_k}{N}. \tag{40}$$

And, define a new variable,

$$k \equiv 1 + \frac{N_k}{N}, \quad e^{\sigma^2} - 1 = \frac{N_k}{N} \tag{41}$$

that is roughly inversely proportional to N over the domain $k \in [1, \infty)$. Which when combined with equation (39) results in,

$$\sigma^2 = \ln(k). \tag{42}$$

Defining a new variable,

$$z \equiv \frac{\sigma^2}{c^2} e^{-\frac{\sigma^2}{c^2}}, \tag{43}$$

and substituting back into equation (31), results in,

$$s = \frac{1}{2}\ln(2\pi e c^2) + \ln(z) + c\ln(\langle u\rangle). \tag{44}$$

If we expand equation (44), to a multivariate case,

$$s = s_0 + a\ln(z) + b\ln(\langle u\rangle) + c\ln(\lambda), \tag{45}$$

we now have the familiar form suitable for regression. Equation (45) was applied to the Bitcoin transaction distribution from the genesis block to early March 2014 [28]. This modeling resulted in an equation of state for the Bitcoin network of,

$$s = 7.9 + 0.61\ln(z) + 0.126\ln(\langle u\rangle) + 0.688\ln(\lambda), \tag{46}$$

where $\langle u\rangle = \lambda\,\mathbb{E}[\mathtt{tx}_{\mathtt{out}}]$, $\mathtt{tx}_{\mathtt{out}}$ are the set of transactions in the studied blocks, the blockchain was binned into bins of 6 blocks of transactions and λ is the total network hash for the block divided by the bin's total block reward, the direct on-chain measure of the marginal utility of bitcoin.

Following the model of modeling the Bitcoin network, write equation (45) in a slightly different form,

$$s = s_0 + a\ln(z) + c\ln(\langle u\rangle) + R\ln(m), \tag{47}$$

where $\langle u\rangle$ is the average income (in absolute terms), m is the specific money supply. Which, when expanded from the individual to the society results in,

$$S = Ns_0 + N\,a\ln\left(\frac{z}{z_0}\right) + N\,c\ln\left(\frac{U}{U_0}\right) + N\,R\ln\left(\frac{M}{M_0}\right) - N(c+R)\ln\left(\frac{N}{N_0}\right), \tag{48}$$

$$U = U_0 e^{\frac{s-s_0}{c}} \left(\frac{z}{z_0}\right)^{-\frac{a}{c}} \left(\frac{M}{M_0}\right)^{-\frac{R}{c}} \left(\frac{N}{N_0}\right)^{1+\frac{R}{c}}. \tag{49}$$

Taking some useful derivatives,

$$\frac{1}{T} \equiv \frac{\partial S}{\partial U} = \frac{Nc}{U} = \frac{c}{\langle u \rangle} \therefore \langle u \rangle = cT, \tag{50}$$

$$\lambda \equiv \frac{\partial U}{\partial M} = \frac{RU}{cM}. \tag{51}$$

Combining equations (50) and (51) results in the familiar,

$$\lambda M = NRT \tag{52}$$

We can also derive the equally familiar polytropic relationship,

$$\lambda M^{\gamma} = Const. \tag{53}$$

Where, for an isentropic process, $\gamma = 1 + \frac{R}{c}$. Continuing with the derivatives,

$$\mu \equiv \frac{\partial U}{\partial N} = \frac{\gamma U}{N} = (c + R)T \tag{54}$$

An additional item of note is that the entropy of the log-normal thermal distribution is maximized when $\sigma = 1$. When applied to income distributions, this results in a $Gini \approx 0.531$. Thus, when we are looking to see what policies promote the greatest social entropy, deviations from a $\sigma = 1$ tell us if the policies are being too restrictive or too permissive. As $\sigma \to 0, Gini \to 0$, and $N \gg N_k$ implying that the population is much greater than the carrying capacity and is too restrictive. Conversely as $\sigma \to \infty, Gini \to 1$, and $N \ll N_k$ implying that the population is too diffuse and lacks adequate structure.

Returning to the Bitcoin modeling and the ideal money law, equation (52), we see the policy impacts of quantitative easing. Expanding the money supply is not stimulative. It is extractive due to equation (54). In modeling a highly simplified Brayton cycle the expansion of the hot combustion gases is approximated by assuming an ideal, isentropic expansion. That represents the maximum useful work extracted from the system. The problem here is that the system is a system of people not a gas.

5. Conclusions

The exclusion of the formal aggregation of microeconomic principles into macroeconomic equations of state has significantly negatively impacted the social sciences. The absence of formality has led to using a self-contradicting normative basis for public policy, Rawls' theory of distributive justice, economic policies that focus on destroying social complexity in promoting some idea of arbitrary equality, because "it feels more right", lead to economic policies that strip wealth from the society to benefit a few, and many other problems.

It is only by formally accounting for entropy in the social sciences that those "sciences" have any hope of approximating a science. While they have made some genuine discoveries, they are mired in feelings and egalitarian ideas that fall apart the moment they are exposed to reality.

If we are hoping to develop policies that promote human welfare, we should look at removing unnecessary restriction and/or providing adequate restriction on their action and then increasing their access to energy. They will take care of the rest on their own.

Supplementary Materials: The following are available online at https://github.com/crabel99/incomeDist: All source code, input data used, and output data generated along with all graphics included in this document.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Gibbs, J.W. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*; Publisher: Charles Scribner's Sons New York, United States, 1902.
2. Jaynes, E.T. Information Theory and Statistical Mechanics. *The Physical Review* **1957** 106 (4), pp. 620-630.
3. von Mises, L. *Human Action: A Treatise on Economics*, 4th ed.; Publisher: Liberty Fund Indianapolis, IN, United States, 2007.
4. von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*, 60th anniversary ed.; Publisher: Princeton University Press Princeton, NJ, United States, 2004.
5. Pfanzagl, J. *Theory of Measurement*; Publisher: John Wiley & Sons New York, Unites States, 1968.
6. Ellsberg, D. Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* **1961** 75 (4), pp.643-669.
7. Allais, M. The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School. In *Expected Utility and the Allais Paradox*, Allais, M.; Hagen, O., Eds.; Publisher: D. Reidel Boston, United States, 1979; pp. 27-145.
8. Yang, J.; Qiu, W. Normalized Expected Utility-Entropy Measure of Risk. *Entropy* **(2014)** 16(7), pp. 3590-3604. https://doi.org/10.3390/e16073590
9. Savage, L.J. *The Foundations of Statistics*, 2nd ed.; Publisher: Dover New York, United States, 1972.
10. Matsoukas, T. *Generalized Statistical Thermodynamics: Thermodynamics of Probability Distributions and Stochastic Processes*; Publisher: Springer, 2018. https://doi.org/10.1007/978-3-030-04149-6
11. Rawls, J. *A Theory of Justice [Kindle Edition]*, Original ed.; Retrieved from Amazon.com, 1971.
12. Smith, A. *An Inquiry into the Nature and Causes of the Wealth of Nations [Kindle Version]*; Retrieved from Amazon.com, 1776.
13. Banerjee, A.; Yakovenko, V.M. Universal patterns of inequality. *New Journal of Physics* **2010**, 12, pp. 1-25.
14. Bell, G. I.; Glasstone, S. *Nuclear Reactor Theory*; Publisher: Van Norstrand Reinhold Company New York, United States, 1970.
15. Duderstadt, J.J.; Hamilton, L.J. *Nuclear Reactor Analysis*; Publisher: John Wiley & Sons, New York, United States, 1976.
16. SOI Tax Stats – Individual Income Tax Returns Publication 1304 (Complete Report). Available online: https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-returns-publication-1304-complete-report (accessed on 29 September 2020).
17. Cavanaugh, J.E. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistics & Probability Letters* **1997**, *33 (2)*, pp. 201-208.
18. Ayers, R.; Warr, B. *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*; Publisher: Edward Elgar Northhampton, MA, United States, 2009.
19. Callen, H.B. *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed.; Publisher: John Wiley & Sons New York, United States, 1985; p. 68.
20. Bitcoin (BTC) Ranks as 6th Largest Global Currency. Available online: https://www.altcoinbuzz.io/cryptocurrency-news/bitcoin-btc-ranks-as-6th-largest-global-currency/ (accessed on 29 September 2020).
21. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 29 September 2020).
22. Bitcoin Hashrate historical chart. Available online: https://bitinfocharts.com/comparison/bitcoin-hashrate.html (accessed 29 September 2020).
23. Annual Energy Review 2011. Available online: https://www.eia.gov/totalenergy/data/annual/pdf/aer.pdf (accessed 29 September 2020).
24. Gross Domestic Product: Implicit Price Deflator. Available online: https://fred.stlouisfed.org/series/GDPDEF/ (accessed on 29 September 2020).
25. Consumer Price Index (CPI) Databases. Available online: https://www.bls.gov/cpi/data.htm (accessed on 29 September 2020).
26. Clementi, F.; Gallegati, M. Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States. In *Econophysics of Wealth Distributions*, Chatterjee, A., Yarlagadda, S., Eds.; Publisher: Springer Milano, Italy, 2005; Volume 3, pp. 3-14. https://doi.org/10.1007/88-470-0389-X_1
27. Landau, L.D.; Lifshitz, E.M. *Statistical Physics*, 3rd ed. Part 1; Publisher: Elsevier New York, United States, 1980; pp. 7-9.
28. Quantifying the Value of Bitcoin. Available online: https://statisticaleconomics.org/2014/03/13/quantifying-the-value-of-bitcoin-2/ (accessed on 29 September 2020).