Article

Elementary Principles in Statistical Economics

Cal Abel [1,*]

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

Abstract: The purpose of this paper is to provide the necessary mathematical justification for applying the methods of econophysics into economics, particularly with introducing the second law of thermodynamics as a fundamental economic constraint. This was done by beginning with the axioms defining game theory and presenting a new set of desiderata that define logical consistency instead of economic rationality. The focus on utility was entirely removed from the derivation of subjective probability. Utility was then reintegrated with the derivation of the entropy functional of the canonical ensemble for the individual. The individual's entropy functional was then aggregated to create a group entropy functional. This approach formally resolved the Allais paradox providing a formal methodology for choice under uncertainty. Because entropy is simply a measure of information, it should only be natural to consider this a fundamental part of economic theory. Macroeconomic models have no formal inclusion of entropy. Because of this, those models ignore the simple fact that the economy is made up of people. As entropy is a direct result of the complexity of human action, it makes little sense for the study of human action not to explicitly use and rely upon entropy.

Keywords: statistical mechanics; information theory; game theory; subjective expected utility; entropy; distributive justice

---

1. Introduction

Physicists tend to be more concerned about explaining what is observed than providing a cohesive underlying theory explaining why things are a particular way. What comes first is the quantification of the observation and then, eventually (although not always), some advance in providing an underlying theoretical basis for the model. This mindset has its roots in the origin of the modern physical sciences. In fact, the development of the study of thermodynamics was driven entirely by the economic need of improving heat engines at the beginning of the industrial revolution [1]. Similarly, the modern chemical industry grew from satisfying the economic needs of improving textile dyeing and petroleum processing [1, 2]. In the physical sciences, theory is always derived from empirical observation.

Given this background, it is understandable that physicists went off and founded the field of econophysics without having any formal theoretical basis in economics – relying entirely on heuristics and intuition to advance their field [3]. In econophysics, they grafted the same mathematical approaches that were developed and refined in physics over the past 200-years into economics [3]. Because there was not a provided theoretical basis with a suitable mathematical explanation, the economic journals have refused publication limiting the acceptance of econophysics in economics [3].

The purpose of this paper is to provide a theoretical basis to formally bridge the work of econophysics into macroeconomics and microeconomics. If history is to be our guide, we need to turn to how thermodynamics became explained entirely by statistical mechanics. This was done in 1902 by Josiah Willard Gibbs in his seminal work *Elementary Principles in Statistical Mechanics developed with especial reference to the rational foundation of thermodynamics* [4]. Gibbs applied Hamiltonian mechanics to create a distribution for a set of particles' distribution-in-phase. Then, he applied information theory[1] to that distribution to derive an expression for the entropy of the system [4] 46-years before Shannon developed information theory [5].

To be an effective bridge, this paper needs to be firmly rooted in the underlying mechanics of microeconomic theory and then to develop from that foundation the necessary structure to provide a rational foundation to econophysics.[2,3] To provide the economic foundation, the paper will use the subjective probability of the individual which was developed as part of subjective expected utility theory, SEU [6, 7]. Because of the shortcomings of SEU [8, 9] and the limitations of economic rationality [10], this paper will have to make some adjustments to SEU to make it work for what we need of it.

The paper will then derive the entropy of an individual using the maximum entropy approach of Jaynes [11] for the discrete case and the generalization of Matsoukas [12] for the continuous case. It is at this point that the approach presented here differs from Gibbs [4] in two ways. Frist, Gibbs considered the aggregate probability of the canonical ensemble from the beginning of his book, we are starting with a similar expression of the canonical ensemble of the individual first and will then aggregate into the group. Second, Gibbs began with continuous distribution functions, while here we begin with discrete distributions later generalizing into continuous distributions. We will then apply some simplifying maximum entropy assumptions taken from Rawls political philosophy [13] to provide a maximum entropy representation of social entropy.

The derived entropy functional will then be extended further into economics to provide the fundamental relationships of econophysics, provide the economic canonical ensemble, derive the various thermodynamic relations, provide an absolute metric of economic utility, and resolve the Allais paradox by presenting a model for choice under uncertainty that incorporates entropy.

The derivation of the functional relations similar to those of thermodynamics allows econophysics to apply their models directly into economics. We hope that by providing a general derivation based on the axioms underlying game theory, that econophysics can gain more purchase in the realm of mainstream economic thought.

2. Foundation

Because the approach taken here relies extensively on information theory, we need to turn to an existing economic theory which formally defines a measure space suitable for stochastic analysis, game theory. Specifically, we will focus on SEU at the point of development as it was in the early 1960's.[4] By "rolling back" to this earlier version of game theory, we will be removing a number of "patches" that have been implemented since then and we will have to spend some time on patching the theory ourselves to make it functional.

This section will present the necessary adjustments to the theory to make it a suitable foundation for statistical economics. These changes are broken down into four main areas: exploring the Allais

---

[1] Gibbs called this his index of probability [4]. Shannon would later call the same thing uncertainty [5]. The expectations of these measures are Clausius' entropy and information entropy respectively. The difference being the application of the Boltzmann constant k to give the Clausius entropy [4].

[2] Because this paper so closely mirrors the approach of Gibbs [4], the author felt that the title should reflect Gibbs' in honor of and acknowledgement to his work.

[3] Gibbs use of the word rationality was based on the observability of a system. In economics, rationality is defined as a being conforming to a narrow set of rules. These rules restrict the domain of what is and is not acceptable economic behavior. For the purposes of this paper, rationality is defined as being that activity which is observable. Because this has to be contrasted with economic rationality, economic rationality will be referred to as "economic rationality" to delineate it from rationality.

[4] There has been nearly 60-years of work on game theory since the early 1960's. The intent with selecting this time of development was to limit the amount of work needed to reconcile later advances in the field and create as manageable of an approach as possible by selecting only what was absolutely needed from the first 20-years of game theory's existence.

paradox, shifting economic rationality to being the much weaker condition of observability, providing the desiderata necessary for consistent reasoning, and then deriving subjective probability and subjective utility under the new framework.

Before beginning in earnest, it is important to go through the pertinent developments in the first 20-years of game theory. The field was launched in 1944 with the release of John von Neumann's and Oskar Morgenstern's book [14].[5] In it, they defined a set of axioms under an ergodic framework to develop a formal measure of a person's utility under a finite set of choices. A decade later, Savage separated the subjective probability from the subjective utility [6] which created SEU as a field of study. As Savage was developing SEU, Allais and Ellsberg among others noted contradictions within the theory and where the theory did not check with experiment [8, 9]. Finally in 1968, Pfanzagl showed how to remove the dependence on ergodicity from subjective probability [7], placing it squarely within the Bayesian framework.

2.1. Paradox

Savage's and Pfanzagl's approaches to subjective expected utility both ran into paradoxes, notably the Allais Paradox [9]. Allais provided a simple game that showed that people would not make decisions that maximized their subjective expected utility (payout) and be consistent with the constraints of economic rationality. Ellsberg presented a similar argument but in a game that presents the paradox more clearly [8], Pfanzagl specifically notes this [7] (p. 206), but fails to reconcile it.

Ellsberg's experiment is represented in Table 1 [8] (pp. 653-654),. In it, there are three different colored balls: red, black, and yellow. There are a total of 90 balls within the urn with 30 of them being red and the remaining 60 balls some unknown distribution between black and yellow. The example presents two different games with the purpose of presenting a contradiction in Savage's rationality. In each of the two games, the outcome of the SEU is the same for each of the wagers, 33.3 and 66.6 for games A and B respectively. What Ellsberg found through experiment is that the individuals express a preference that is not consistent with the lack of illusion principle, "virtually identical wagers presented in different ways have the same utility" [7] (p. 202).

Under the lack of illusion principle, wagers I and III are virtually identical because in each the red balls are selected. Wagers II and IV are identical because of the preference for the set of black/yellow balls. What Ellsberg found was that during experiments that the outcomes were not consistent with the definition of economic rationality. 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 [8]. Ellsberg [8] (p. 655) and Pfanzagl [7] (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 that the number of black balls, α, was uniformly distributed on $\alpha \in [0,60]$. If an individual had a different prior probability other than the uniform distribution for the distribution of α, that would affect their selection of the different outcomes and affect their expectations, changing the game in their minds.

The entropy in Table 1 is calculated by using Shannon's formulation of information entropy:

$$H(p_1 \cdots p_n) = -K \sum_i p_i \ln p_i \quad [11] \text{ (p. 622)},^6 \tag{1}$$

where p_i is the probability of the i^{th} event, and K is a positive constant which we will assume $K = 1$.

[5] Utility as it is discussed in game theory generally refers to the measure of the reward of the payout from a wager made in a game.

[6] Referring to Shannon's entropy as H is done here because this is the typical variable in information theory for referring to information entropy. This will be the only time in this paper that entropy will be denoted as H. For the remainder of the paper entropy will follow the convention of thermodynamics and be either a lower case s for the specific entropy of the individual or the upper case S for the entropy of the group.

Table 1. Summary of Ellsberg's Game [8] (pp. 653-654) where $\alpha \in [0,60]$ is an unknown number to the player. The highlighted wagers demonstrate those that were empirically preferred in each game.

	Odds	30	α	$60-\alpha$		
		Payout				
Game	Wager	red	black	yellow	Expected Payout	Entropy
A	I	100	0	0	$100 \cdot \frac{1}{3}$	0.637
	II	0	100	0	$100 \cdot \frac{1}{3}$	0.637
B	III	100	0	100	$100 \cdot \frac{2}{3}$	0.637
	IV	0	100	100	$100 \cdot \frac{2}{3}$	0.637

In these games, A and B, the information content of each of the pairs of choices is equivalent, as are the expected payouts. The outcomes of either of the games are indistinguishable from each other. Where the paradox arises is that the preference for attributes that do not affect the underlying game. Like having a preference for either being the shoe, the dog, or the battleship in Monopoly. For the purpose of the game the set of choices, (I,III), (I,IV), (II,III), (II,IV) all represent identical and indistinguishable outcomes. If this represents a contradiction to the condition of rationality, then the condition of rationality is too strong. Saying that one is acting irrationally for choosing the dog as their playing piece when they play with their family, but then choose the battleship when playing with friends is an equivalent statement to the irrationality described in the paradox. Ultimately the choice is the same, they are choosing to play the game.

The Allais paradox is created entirely by the constraints of Savage's theory, particularly from the narrowness of the definition of economic rationality. *The theory must admit the set of all possible outcomes*. We cannot arbitrarily cull what, through observation, are clearly allowed states. Nor can we distinguish between states of equivalent outcome even if they are presented differently. For these reasons, we have to abandon the conventional formulation of rationality and develop a new one that is consistent with *all of our observations*.

2.2. Rationality

One of the areas where Savage and Pfanzagl shared their approach is with the "lack of illusion" principle and the "sure-thing" principle which they used to define rational economic behavior. Pfanzagl defined them as:

> **"Lack of Illusion" Principle:**
> Virtually identical wagers presented in different ways have the same utility.
> **"Sure-thing" Principle:**
> i. The utility of a wager remains unchanged if an outcome is substituted by an outcome of the same utility.
> ii. If one of the outcomes is substituted by an outcome with higher utility, then the utility of the wager either increases or remains unchanged for all possible substitutions of this outcome. [7] (p. 202)

von Mises notes that "[h]uman action is necessarily always rational" [10] (p. 19). He is referring to the idea that for any individual, their choices are consistent within their own frame and judgment and are then by definition rational. And that by applying an arbitrary set of judgments on rationality, we exclude whole sets of individual choice from economic consideration. Thus, we need to rethink defining rational behavior as the "sure-thing principle and the lack of illusion principle" [7] (p. 205). Instead, we will use the definition of rationality as being that action which is observable – fully adopting the rationality principle of von Mises [10] (p. 12). Furthermore, by restricting consideration of the theory only to observable events of human action, we have also adopted von Mises action axiom [10] (p. 11).

Clearly, we need a different set of principles, desiderata, that seek to impose constraints of logical consistency instead of forcing an arbitrary definition of rationality.

2.3. Desiderata

The problem is that if we are going to follow von Mises principle that all action be rational, we need to take a more general approach. Where instead of demanding rationality, we demand logical consistency. For these principles we turn to Jaynes' desiderata [15] (pp. 17-19). The desiderata below are modified slightly from what Jaynes originally wrote with the intention of improving clarity not modification of the meaning.

1. Degrees of plausibility are represented by real numbers.
2. Qualitative correspondence with common sense.
3. Consistency
    a. If a conclusion can be reasoned out in more than one way, then every possible way must lead to the same result.
    b. All of the evidence that is relevant to a question is always taken into account. Some of the relevant information is not arbitrarily ignored, with conclusions based only on the remaining restricted information. In other words, reasoning is done in a completely nonideological manner.
    c. Equivalent states of knowledge are represented by equivalent plausibility assignments. That is, if in two problems the state of knowledge is the same (except perhaps for labeling the propositions), then the same plausibilities are assigned to both.

It is clear to see that desiderata 3.a is a reformulation of the "lack of illusion" principle, desiderata 3.b is an equivalent way of stating von Mises' condition of rationality and the action axiom, and that desiderata 3.c is a reformulation of the first half of the "sure-thing" principle. The second half of the sure thing principle is contained in desiderata 2. To demonstrate this, we will use Jaynes' example of how changing the given information changes the plausibility of an event.

Let the plausibility of a number of subsets of events of all of the available events $\mathcal{E}$ be represented as being $A, B, \text{and } C$, where $A, B, C \in \mathcal{E}$ and the plausibility of the events being represented as $(A), (B), \text{and } (C)$. We can describe the plausibility of event A given event C as being $(A|C)$ and similarly for event B as $(B|C)$.

If we have a new event C' that increases the plausibility of event A, we can represent that as

$$(A|C') > (A|C); \tag{2}$$

but the updated information did not change the plausibility of B given A:

$$(B|AC') = (B|AC). \tag{3}$$

This can only increase, never decrease the plausibility of the truth of events A and B:

$$(AB|C') \geq (AB|C) \text{ [15] (p. 18).} \tag{4}$$

Savage and Pfanzagl were right in seeing the need for need of the sure thing principle and the lack of illusion principle as providing a measure of consistency. However, how they applied it forced too strong of a condition. The correct measure of consistency is applied through the desiderata of this section, which also reconcile the philosophical basis of the Austrian school, von Mises, with other theories of economics such as game theory and, as we will show later, macroeconomics. We now have a sufficient basis to derive subjective probability under the new desiderata independent of the derivation of economic utility.

2.4. Subjective Probability and Utility

Because of the length of the proofs of the derivations of subjective probability in Pfanzagl [5] and in Jaynes [15], an interested reader is recommended to those references for more detail. The "lack of illusion" and the "sure-thing" principles that Pfanzagl specified are not axioms. He used them as tools, desiderata, to guide the selection of axioms. By changing the focus of what is measured from utility to plausibility. By adopting Jaynes desiderata, we have shifted the focus of game theory entirely to the plausibility of a set of individual choices.

We will use Pfanzagl to develop the event space. Then we will use Jaynes to derive probability theory. Finally, we will adopt a set of axioms to derive the utility functional. This approach is similar to Savage's where he derived subjective probability first and utility second [6]. Because of the difficulty in finding [7] the pertinent parts of that work relevant for the derivation are included in Appendix A. Those portions will be referenced as "A.x.x.x" where "x.x.x" is the relevant section number in Appendix A from the original text.

There are two main sets used in deriving game theory, events and outcomes. The set of all events is $\mathcal{E}$ and the set of possible outcomes is A. The event is tied to some uncertain process. The outcome can be thought of as the wager amount depending on whether a particular event obtains or does not obtain. The event space is defined in A.12.2.1 as being a Boolean algebra. It relies on all of A.12.1 to provide the formal relationships. The axioms defining the system of events are A.12.1.1 (the definition of a Boolean algebra). This approach is consistent with the Kolmogorov system of probability, which Jaynes showed as being contained within his derivation of probability theory [15] (Chapter 2 and Appendix A).[7]

With probability as being a property of the event set, and the outcome space being "different quantities of a commodity" [7] (p. 202), it is clear that utility ties into the probability measure through the event space. Pfanzagl didn't use the sure thing principle or the lack of illusion principle until after he had defined the event space. Because the desiderata adopted here transform the economic rationality of the sure thing principle and the lack of illusion principle into a constraint of logical consistency on the plausibility of events, the conditions needed to develop utility are now contained in the probability theory, specifically, A.12.2.12, A.12.3.2, and A.12.3.3.

To formally show how the desiderata subsume economic rationality is beyond the need of this paper. Without assuming the economic rationality constraints, we can take A.12.2.7, A.12.2.12, A.12.3.2, and A12.3.3 as the axioms that define the utility space [7] (p.207). These axioms are used to derive the utility functional [7] (Sections 12.4 and 12.5) and are appropriate constraints on utility space, but are not appropriate for the reasoning space, there we have to apply Jaynes' desiderata, which is why the proof was split as such.

At this point subjective probability, as modified here, is sufficient to develop an expression for the specific entropy of an individual. It is important to note that little has changed in the formulation of game theory other than shifting the consistency condition from the utility to that of reasoning. The next section will derive the entropy functional.

3. Derivation

It is impossible to not emphasis enough how crucial Jaynes' collected works are in generalizing statistical mechanics. He saw Gibbs mathematical foundation of thermodynamics as being an application of information theory to what Gibbs called the "distribution-in-phase". Gibbs started his derivation with the Hamiltonian and then derived the canonical distribution. For him, the mechanics were known, but the distribution was not. In economics, we have the opposite problem: the distribution is known, but not the mechanics.

Because Jaynes saw Gibbs so clearly [11], he cut straight to the essence of the problem – the entropy functional is a consequence of the distribution. This is why the desiderata from the previous section are so critical to the foundation of the overall theory. If any relevant observations are excluded, the distribution is affected, which in turn affects the entropy functional. Thus, by excluding relevant information it becomes impossible for us to describe what is.

The derivation of this section contains two main parts: the derivation of the entropy functional for the individual and the derivation of the entropy functional for the group. The derivation of the individual entropy functional will follow two paths, first the discrete path following Jaynes method

[7] Jaynes showed that the axioms Kolmogorov used were a consequence of Jaynes' desiderata.

[11] and second the continuous generalization following Jaynes [16] and Matsoukas [12]. The aggregation of the group will require a turn to political philosophy [13] for guidance.

3.1. Individual Entropy

3.1.1. Discrete Case

By adopting the modified subjective probability, we can express our knowledge (represented by probability measures) of the observable outcomes from a finite set of options, with no knowledge of the actual underlying mechanics, the utility function. At this point, our knowledge of the individual's choices can be expressed by the subjective probability for a particular event, p_i, where $p_i \in P\ \forall\ i \in [1,n]: P \in \mathcal{E}$, $\mathcal{E}$ is the set of possible events, and where the subjective probability has the normalization constraint, $\sum_i p_i = 1$. We let x be a variable that assumes discrete values $x_i\ \forall\ i \in [1,n]$ corresponding with p_i. We also define an arbitrary mapping $u{:}\,x \longrightarrow u(x)\ \forall\ x_i \in x$ that satisfies Pfanzagl's four axioms that derive utility. The general method of the proof is to maximize equation (1) through the use of Lagrangian multipliers. Because of Jaynes' [11] (pp. 622-623) proof's simplicity it will be replicated here in a slightly more condensed form, and restrict ourselves to the canonical representation.

Because the entropy of equation (1) is solely a function of the probability p_i, reformulate the probability to include the Lagrangian multipliers ψ and β.

$$p_i = e^{\psi\varphi_0 - \beta u(x_i)}, \tag{5}$$

subject to the constraints,

$$\mathbb{E}[\mathrm{u}(x)] = \sum_i p_i u(x_i) = \langle u \rangle. \tag{6}$$

The expectation of the 0th moment, $\varphi_0 \equiv 1$, must be $\mathbb{E}[\varphi_0] = 1$. This is Jaynes' normalization constraint [11] (p. 622). It is a more general statement of Gibbs' *conservation of the extension-in-phase* [4] (p. 10). The normalization constraint directly results in the partition function,

$$Z[\beta] = e^{-\psi} = \sum_i e^{-\beta u(x_i)}. \tag{7}$$

The remaining Lagrangian multiplier can be determined from,

$$\mathbb{E}[u(x)] = -\frac{\partial}{\partial\beta}\ln(Z[\beta]) = -\frac{\partial\psi}{\partial\beta}. \tag{8}$$

Borrowing from Gibbs,[8] we define the uncertainty as being

$$\eta_i = \ln p_i = \psi - \beta u(x_i). \tag{9}$$

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

$$s = -\sum_i p_i\eta_i = -\psi + \beta\langle u \rangle. \tag{10}$$

Rearranging and defining $\psi \equiv \beta f$,

$$f = \langle u \rangle - \frac{s}{\beta}. \tag{11}$$

[8] The remaining derivation will follow the form Gibbs used for the continuous case in his derivation of the second law of thermodynamics [4] (pp. 42-44). In following section, it will be sufficient to show the proper form of the entropy functional for the continuous case without having to duplicate the derivation of the second law as the methodology there follows Gibbs almost exactly.

Assuming the utility depends not only on the individual, but also upon a set of external variables that they are not regarded as forming any part of the individual, although their values affect the individual. We can express their effect on the utility by,

$$\frac{du}{da_k} = \varphi_k(x_i). \tag{12}$$

Taking the total differential of the partition function results in,

$$dZ[\beta] = e^{-\beta f}(-\beta df - f d\beta) = -d\beta \sum\nolimits_i u(x_i)e^{-\beta u(x_i)} - \sum\nolimits_k \beta da_k \sum\nolimits_i \frac{du}{da_k} e^{-\beta u(x_i)}. \tag{13}$$

Multiplying equation (13) by, $\frac{-1}{\beta}e^{\beta f}$ and substituting in equation (12) provides,

$$df + \frac{f}{\beta}d\beta = \frac{1}{\beta}d\beta \sum\nolimits_i u(x_i)e^{\psi-\beta u(x_i)} + \sum\nolimits_k da_k \sum\nolimits_i \varphi_k(x_i)e^{\psi-\beta u(x_i)}. \tag{14}$$

Substituting equation (11) and resolving the expectations results in and substituting $\beta \equiv \frac{1}{T}$

$$df = -sdT + \sum\nolimits_k \langle\varphi_k\rangle da_k. \tag{15}$$

Taking the total differential of (11) results in,

$$df = d\langle u\rangle - Tds - sdT. \tag{16}$$

And substituting the result with equation (15) provides,

$$d\langle u\rangle = Tds + \sum\nolimits_k \langle\varphi_k\rangle da_k. \tag{17}$$

So far, there has been no loss in generality in deriving equation (21) which is the mathematical expression of the second law of thermodynamics for a reversible process [4] (p. 44). However, equation (17) is only applicable for a finite set of discrete events described by a probability mass function. It is not applicable to continuous density functions. To derive a relationship for a continuous probability density function, we need to turn our attention to Jaynes' Brandeis lectures [16].

3.1.2. Continuous Case

Jaynes [16] derived the entropy functional of a continuous density function, taken as the continuous limit of the discrete case, to be

$$\mathrm{S[h]} = -\int dx\, h(x) \ln\frac{h(x)}{m(x)}, \tag{18}$$

where $h(x)$ is a continuous density function for all x and where $m(x)$ is the invariant "measure" of $h(x)$ [16] (p. 202).[9] Which, if the distribution h is properly normalized, $\int dx\, h(x) = 1$, equation (18) reduces to the familiar Shannon differential entropy functional,

$$\mathrm{S[h]} = -\int dx\, h(x) \ln h(x) \;\; [12]. \tag{19}$$

Using a similar method as in the discrete case, we will maximize the entropy functional, equation (19) subject to a set of constraints, but use the grand canonical ensemble. Define the probability density function to be estimated as,

$$f(x) = e^{\psi\varphi_0(x)-\beta u(x)+\beta\sum_k a_k\varphi_k(x)}, \tag{20}$$

where a_k and $\varphi_k(x)$ are defined as before with $\varphi_0(x) = 1$ and the constraints:

$$\mathbb{E}[\varphi_k\,(x)] = \int dx\, \varphi_k(x)f(x) = \langle\varphi_k\rangle. \tag{21}$$

[9] Matsoukas [12] (p. 205 and pp. 221-225) shows how to satisfy Jaynes' invariance condition and an interested reader is referred there for further detail.

To keep $f(x)$ properly normalized, $\mathbb{E}[\varphi_0\,(x)] = \int dx\, f(x) = 1$, allowing the use of the entropy functional equation (19). This constraint directly results in the partition function,

$$Z[\beta, a_k] = e^{\psi} = \int dx\, e^{-\beta u(x)+\beta \sum_k a_k \varphi_k(x)}. \tag{22}$$

The remaining Lagrangian multipliers can be obtained from,

$$\beta\mathbb{E}[\varphi_k(x)] = -\frac{\partial}{\partial a_k}\ln(\mathrm{Z}) = -\frac{\partial \psi}{\partial a_k}. \tag{23}$$

And the resulting maximum entropy estimate of the individual's demonstrated preference for a continuous distribution is,

$$s = \beta\langle u\rangle - \sum\nolimits_k \beta\langle\varphi_k\rangle a_k. \tag{24}$$

3.2. Group Entropy

Before proceeding with the group aggregation, we need to assume logical independence of each individual's action. This assumption has a physical meaning of the freedom of individual choice – there is no mind control or forced coercion (a truly ideal representation of the human action). 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" [13] (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 [10]. As Jaynes notes, the assumption of the principle of maximum entropy is Laplace's "principle of insufficient reason" [11] (p. 622).

Mathematically this results in the group's density function,[10] $F[\boldsymbol{X}]$, being separable in $x_j\ \forall\, x_j \in \boldsymbol{X}$,

$$F[\boldsymbol{X}] = \prod_j f_j(x_j), \tag{25}$$

where x_j is the domain of j^{th} individual's action, $\boldsymbol{X}$ is the vector comprised of the x_j's, and f_j is the density function for the of j^{th} individual in the group. From equation (19),

$$\mathrm{S}[F] = \int d\boldsymbol{X}\, F[\boldsymbol{X}]\ln F[\boldsymbol{X}]. \tag{26}$$

Due to the logical independence of the actors, equation (26) reduces to,

$$S = \sum_j s_j. \tag{27}$$

At this point, no assumptions have been made outside of the axioms used to derive subjective probability other than the maximum entropy condition of each individual's expression of their freedom to choose. Thus, equation (27) formally represents a *complete* group entropy functional. 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" [13] (Ch.1 Sect. 3). The veil of ignorance 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,

$$S = Ns, \tag{28}$$

[10] Because of the high dimensionality of human action, we will use the continuous probability to define our knowledge of each person's action.

where, N is the integer value of 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 an invisible hand [entropy] to promote an end which was no part of [our] intention" [17] (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 societal level through debate. The justification of reducing liberty on the small to promote overall liberty is Rawls' first priority rule [13] (Ch. V Sect. 46) and is warranted.

4. Extension and Consequences

So far, the work presented here has been abstract, proving the entropy functional of stochastic models that satisfy the axioms of a Boolean algebra used to define the set of events in game theory. While this has provided some additional insights, it is not very useful. Our next task is to explore the more traditional relationships.

4.1. The Euler Relation

Combining equations (24) and (28), we have the Euler relation,

$$S = N\left(\beta\langle u\rangle - \sum\nolimits_k \beta\langle\varphi_k\rangle a_k\right) = \beta\langle U\rangle - \sum\nolimits_k \beta\langle\varphi_k\rangle A_k. \tag{29}$$

Where, $\langle U\rangle$ and A_k are the extensive macroeconomic variables, held invariant in the canonical distribution. Because $A_k = \mathrm{N}\, a_k$ and $U = N\, u$, equation (29) is a first order homogenous equation satisfying $S(\lambda\, A_1, \cdots, \lambda\, A_k) = \lambda\, S(A_1, \cdots, A_k)$. Differentiating the homogeneity condition with respect to λ results in,

$$S(A_1, \cdots, A_k) = \sum_k \frac{\partial S(\lambda\, A_1, \cdots, \lambda\, A_t, \cdots)}{\partial(\lambda\, A_k)} \frac{\partial(\lambda\, A_k)}{\partial\lambda} = \sum_k \frac{\partial S(A_1, \cdots, A_t, \cdots)}{\partial A_k} A_k. \tag{30}$$

and the intensive macroeconomic parameters are $\beta = {}^{\partial S}/_{\partial\langle U\rangle} = {}^{\partial s}/_{\partial\langle u\rangle}$ and $\beta\langle\varphi_k\rangle = -\langle {}^{\partial S}/_{\partial A_k}\rangle = \langle {}^{\partial s}/_{\partial a_k}\rangle$.

Writing the first differential of $S = S(A_1, \cdots, A_k)$ as,

$$dS = \frac{\partial S}{\partial\langle U\rangle} d\langle U\rangle + \sum_{j=1}^{k} \left.\langle\frac{\partial S}{\partial A_j}\rangle\right|_{A_1,\cdots,A_k} dA_j, \tag{31}$$

Provides us with the familiar differential form of the fundamental equation of thermodynamics in its entropic representation.

4.2. The Gibbs-Duhem Relation

Solving equation (31) for utility, we have,

$$\langle U \rangle = TS + \sum_k \langle \frac{\partial \langle U \rangle}{\partial A_k} \rangle A_k. \tag{32}$$

Where,

$$\langle \frac{\partial \langle \mathrm{U} \rangle}{\partial A_j} \rangle = -\frac{1}{\beta} \langle \frac{\partial S}{\partial A_j} \rangle \text{ [18] (p. 477).} \tag{33}$$

And a resulting first differential of,

$$d\langle U \rangle = TdS + \sum_{j=2}^{k} \langle \frac{\partial \langle U \rangle}{\partial A_j} \rangle \Big|_{A_1,\cdots,A_k} dA_j \text{ [11]}. \tag{34}$$

[11]. Taking the total differential of equation (32) and subtracting equation (34) results in the Gibbs-Duhem relationship,

$$0 = SdT + \sum_{j=2}^{k} A_j d \langle \frac{\partial \langle U \rangle}{\partial A_j} \rangle \Big|_{A_1,\cdots,A_k}. \tag{35}$$

4.3. The Economic Variables[12]

Along with utility, U, we introduce two new extensive variables, M and N to specify the money supply and the population respectively. We define the intensive parameters of the extensive variables M and N as being the marginal utility of money, λ, and the economic potential of the individual, μ:

$$-\frac{\partial U}{\partial M} \equiv \lambda, \tag{36}$$

$$\frac{\partial U}{\partial N} \equiv \mu. \tag{37}$$

We can look at the other extensive parameters, A_j, having associated intensive parameters of $\frac{\partial U}{\partial A_j} \equiv p_j$. This results in equations (32), (34), and (35) having the respective forms of,

$$U = T\,S - \lambda\,M + \mu\,N + \sum_{j=4}^{k} p_j\,A_j. \tag{38}$$

$$dU = TdS - \lambda dM + \mu dN + \sum_{j=4}^{k} p_j dA_j. \tag{39}$$

$$0 = SdT - Md\lambda + Nd\mu + \sum_{j=4}^{k} A_j dp_j. \tag{40}$$

The sign in equation (36) is there to follow how we think about the value of money in our society. It is analogous to pressure in a physical system and is empirically justified [19].

---

[11] Comparing equation (21) multiplied by N to equation (34) shows the equivalence and compactness of Jaynes method, and serves to prove the aggregate of the discrete case for the group.
[12] At this point we will drop the symbols of the expectations for clarity.

4.4. The Economic Potentials

In general, the main economic potentials are the various partial Legendre transformations of $U[S, M, N, A_4, A_5, \cdots]$ that replace:

- entropy by temperature, Helmholtz potential ($F[T, M, N, A_4, A_5, \cdots]$),
- the money supply by the marginal utility of money, enthalpy ($H[S, \lambda, N, A_4, A_5, \cdots]$),
- simultaneously entropy with temperature and the money supply with the marginal utility of money, Gibbs potential ($G[T, \lambda, N, A_4, A_5, \cdots]$), and
- simultaneously entropy with temperature and the population with the economic potential of the individual, grand canonical potential ($U[T, M, \mu, A_4, A_5, \cdots]$).

There are a number of other potentials that arise from different combinations of partial Legendre transforms of utility. These other partial transforms are unnamed in thermodynamics and only arise infrequently in physics [18] (p. 148). For the application of the transforms and for further discussion a reader is recommended to [18] (pp. 146-148) and other references, e.g. [4, 20], which can provide more detail. We will provide the results of [18] to their economic analogs here.

Helmholtz Potential, $\boldsymbol{F \equiv U[T]}$

$$dF = -SdT - \lambda dM + \mu dN + \sum_{j=4}^{k} p_j dA_j. \tag{41}$$

Enthalpy, $\boldsymbol{H \equiv U[\lambda]}$

$$dH = TdS + Md\lambda + \mu dN + \sum_{j=4}^{k} p_j dA_j. \tag{42}$$

Gibbs Potential, $\boldsymbol{G \equiv U[T, \lambda]}$

$$dG = -SdT + Md\lambda + \mu dN + \sum_{j=4}^{k} p_j dA_j. \tag{43}$$

Grand Canonical Potential, $\boldsymbol{U[T, \mu]}$

$$dU[T, \mu] = -SdT - \lambda dM - Nd\mu + \sum_{j=4}^{k} p_j dA_j. \tag{44}$$

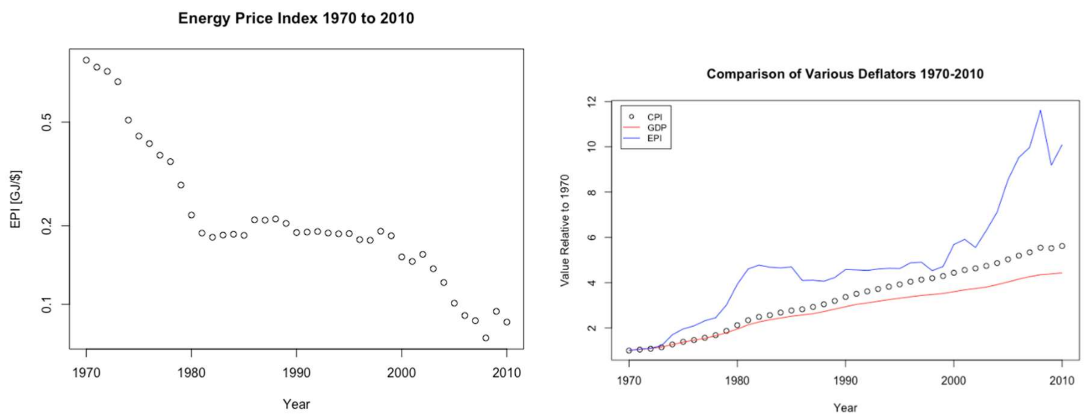
The complete Legendre transformation is represented by the Gibbs-Duhem relationship, equation (40) where $U[T, \lambda, \mu, p_4, p_5, \cdots] = 0$.

4.5. The Marginal Utility of Money

As we have shown, utility is mathematically analogous to internal energy in physics. This, however, is not just an analogy. All human action is constrained by the laws of thermodynamics, thus for us to act, we must expend energy. If our utility is a measure of our ability to act, then the natural unit of utility is energy. Because the metric space of utility is defined as being "different quantities of a simple commodity (e.g. money), filling an interval" [7] (p. 202), energy suffices as a simple commodity filling an interval just as any other arbitrary commodity such as money does. For this reason, we are formally justified in defining the canonical form of utility as energy. For an empirical justification, we need to turn elsewhere.

Ayers and Warr presented an econometric justification for the inclusion of exergy, useful work, into a production function [21]. They found that 80% of GDP can be explained solely by the exergetic input into the economy. By measuring the value of a currency in terms of energy, we have an absolute measure of the currency as energy is a conserved quantity.

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 [22] (Table 1.5). This is from data aggregated from utilities, refineries, and distributors across the nation [22]. Figure 1 shows the estimated Energy Price Index, EPI, from EIA data [22] (Table 1.5), the GDP deflator [23] and the Consumer Price Index- Urban Consumers CPI-U [24].



(a) (b)

Figure 1. (a) The Energy Price Index derived from [22] (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 [24], Gross Domestic Product Deflator [23], and the inverted EPI from (a). All indices are normalized relative to their 1970 values.

4.6. The Approach to Certainty Through Experience

Savage presented an example of how certainty is approached through experience [6] (Section 3.6). It is worth noting here and deriving it under a statistical economics framework using Jaynes' method [11] (p. 625) which is duplicated below using the notation presented in this paper.

We begin by defining x_i as discrete numbers corresponding to events $P_i \in \mathcal{E}$. Let n be a non-negative integer, and let ϵ be a small positive number,

$$x_1^{n+1} = \epsilon, \qquad x_{i+1} - x_i = \frac{\epsilon}{x_i^n} \; \forall \, i = 1, 2, \cdots. \tag{45}$$

According to equation (45), $\lim_{i\to\infty} x_i$ is unbounded. However, the density of the points increases at a rate determined by n. By choosing a small enough ϵ, we can make the density of points high enough in the vicinity of any x allowing the approximation as a continuous function, $f(x)$.

$$\lim_{i\to\infty} \sum_i f(x_i) = \int dx \, f(x)\rho(x). \tag{46}$$

Where from (45), $\rho(x) = x^n/\epsilon$.

Consider the forward problem, given $\langle x \rangle$ provide a maximum entropy estimate of x^2. Using equation (22) with $k = 1$ and $\varphi_1(x) = x$,

$$Z[\beta] = \int dx \, \rho(x) e^{-\beta x} = \frac{n!}{\epsilon \beta^{n+1}}. \tag{47}$$

From equation (23),

$$\mathbb{E}[x] = -\frac{\partial}{\partial \beta} \ln(Z) = \frac{n+1}{\beta}. \tag{48}$$

Then,

$$\langle x^2 \rangle [\langle x \rangle] = Z^{-1} \left[\int dx \, x^2 \rho(x) e^{-\beta x} \right] = \frac{n+2}{n+1} \langle x \rangle^2. \tag{49}$$

For the inverse problem, given $\langle x^2 \rangle$ estimate x. Let $\varphi_1(x) = x^2$,

$$Z[\beta] = \int dx\ \rho(x) e^{-\beta x^2} = \frac{\sqrt{\pi} n!}{2^{n+1} \left(\frac{n}{2}\right)!\, \epsilon \beta^{n+1}}. \tag{50}$$

$$\mathbb{E}[x^2] = -\frac{\partial}{\partial \beta} \ln(Z) = \frac{n+1}{2\beta}. \tag{51}$$

$$\langle x \rangle [\langle x^2 \rangle] = Z^{-1} \left[\iint dx\ x \rho(x) e^{-\beta x^2} \right] = \sqrt{\frac{n+1}{2}} \frac{\left(\frac{n}{2}\right)!}{\left(\frac{n+1}{2}\right)!} \sqrt{\langle x^2 \rangle}. \tag{52}$$

These provide,

$$\mathrm{Var}[\mathrm{x}] \equiv \langle x^2 \rangle - \langle x \rangle^2 = \frac{\langle x \rangle^2}{n+1}\ [11]\ (\text{p. } 625). \tag{53}$$

Thus, as n increases, the variance decreases $\propto 1/(n+1)$. As more information comes in, the subjective definition of the event becomes more pronounced, reducing the number of ways something can be contextualized, e.g. the definition of words, human experience, concepts, etc. For something that is highly corroborated, $n \gg 1$, it becomes near certainty. This reasoning can be used to describe how markets are incredible information aggregators. Furthermore, we can see why economics, particularly macroeconomics, focused so strongly on deterministic models where the uncertainty, entropy, was entirely ignored.

4.7. Choice Under Uncertainty

We now develop the statistical economic foundation for *choice under uncertainty* that incorporates both economic utility and social entropy. Before proceeding we need to consider the impact of selecting the appropriate ensemble by looking at what is constant between the choices. If utility is constant between choices, the microcanonical ensemble is appropriate. If the utility is not constant, then the canonical ensemble is appropriate. The incorporation of entropy into choice has had work done previously [25], it resulted in complicated forms that are not intuitive and seek the introduction of measures, double counting the number of states of a system, etc.

We recall from statistical mechanics that the canonical ensemble,

$$Z[T] = \sum_i e^{-\frac{u(x_i)}{T}}, \tag{54}$$

$$f = -T \ln(Z[T]), \text{ and} \tag{55}$$

$$f = \langle u \rangle - Ts, \tag{56}$$

represents systems in contact with a thermal reservoir. The Helmholtz potential is minimized in statistical equilibrium. However, individuals are trying to maximize their potential compared to each other – playing Maxwell's demon. Which, when played out with a system of demons, will result in a state of maximum entropy.

In the case of Maxwell's demon, the thermal reservoir is the person making the choice between the various wagers. Their temperature represents their preference for risk – marginal utility of entropy – temperature. A low-risk tolerance would be a high temperature a high-risk tolerance would be a low temperature from the sign in equation (56), which shows how the expected utility of the outcome is modified by the uncertainty. In the case of equal utility, the wager with the lowest entropy (less risky option) is the one that is selected. If as discussed previously, the entropies and utilities are equal, the set of choices are indistinguishable from each other and are considered equivalent. If there is a difference in the empirical preference between the choices, then some additional information outside of what is presented is needed to resolve the difference. The canonical approach results in the selection of the greatest potential as a rule for individual choice.

We turn to two examples of presented in [25] (Table 2 and Table 3) presented here as Table 2 and Table 3 respectively. Table 2 represents a comparison between two games where the utilities between the wagers for a particular game are the same but have different entropies. Table 3 presents the case where utility and entropy of the wagers are different within each game.

Table 2. Example of equivalent utilities within games where the shaded wager represents the one that is empirically preferred for a particular game [25] (p. 3598 Table 2).

Game	Wager	Outcomes and Their Corresponding Probabilities		$\langle u \rangle$	s
A	a_1	6000 0.45	0 0.55	2700	0.6881
	a_2	3000 0.90	0 0.10	2700	0.3251
B	b_1	6000 0.001	0 0.999	6	0.0079
	b_2	3000 0.002	0 0.998	6	0.0144

We can see from Table 2 that the preference relation, $\prec$ A.3.2, between the sets of games should be $a_1 \prec a_2$ and $b_2 \prec b_1$ $\forall T$ because $s[a_1] > s[a_2]$ and $s[b_2] > s[b_1]$. This corresponds with the empirical preference [25].

Table 3. Example of varying utilities within games where the shaded wager represents the one that is empirically preferred for a particular game [25] (p. 3599 Table 3).

Game	Wager	Outcomes and Their Corresponding Probabilities		$\langle u \rangle$	s
A	a_1	3000 1		3000	0
	a_2	4000 0.80	0 0.20	3200	0.5
B	b_1	3000 0.25	0 0.75	750	0.56
	b_2	4000 0.20	0 0.80	850	0.5

In Table 3, because the expected utilities are not same between the wagers, we need to consider the risk preference of the individual, Figure 2. We can conclude that in general the risk tolerance of the individuals is relatively low because for game A the empirical preference was $a_2 \prec a_1$ which holds $\forall T > 399.7$. The empirical preference demonstrated in game B, $b_1 \prec b_2$, holds true $\forall T$.

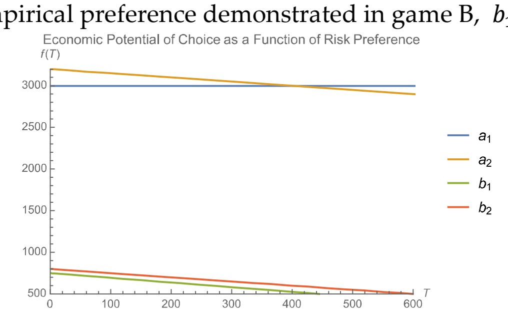


Figure 2. The economic potential of an individual's choice as a function of the individuals risk preference computed from equation (56) and Table 3.

5. Conclusions

This paper began with the axiomatic definitions of game theory and derived subjective probability under a new set of desiderata independent of utility. With this new foundation of microeconomics, the paper then derived the entropy functional as an expression of the Euler relation for the macroeconomy. The concept of economic utility was reintroduced and the associated fundamental relationships (differential, Euler, and Gibbs-Duhem) were derived. The concept of the canonical ensemble for economics was introduced along with the corresponding extensive measures. The basis of economic utility as being energy was proposed. And finally, the Allais paradox was resolved by maximizing the Helmholtz potential for choice under uncertainty.

The full integration of microeconomic principles into macroeconomics introduced the concept of entropy into both micro and macroeconomics. Entropy has to the author's knowledge never been formally integrated into the study of economics. This is especially surprising because entropy can be used as a measure of complexity/uncertainty of a system and entire fields like game theory are focused on the study decisions made under uncertainty. While entropy was acknowledged early in the development of game theory [6] (p. 50), it was left as a curiosity, more of a mathematical footnote than anything else.

From the standpoint of a practitioner of the physical sciences the absence of an entropic term when dealing with an aggregate is a glaring omission. When combined with the intentional exclusion of information (observed events) that don't comport to a specified theory, it is little wonder that economics has been mired in controversy, paradox, and confusion. The lack of acknowledgement of the constraint of the second law of thermodynamics in economics is a scientific failure, *sin qua non*.

"The law that entropy always increases holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations – then so much the worse for Maxwell's equations. If it is found to be contradicted by observation – well, these experimentalists do bungle things sometimes. But if your theory is found to be against the Second Law of Thermodynamics, I can give you no hope; there is nothing for it to collapse in deepest humiliation [26]."

Funding: This research received no external funding.

Acknowledgments: The author is grateful to the patient critique of the reviewers and editors who helped identify deficiencies in the author's work and especially for forcing the author to refine the scope and structure of the article into something far better than the original composition.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix contains the axioms used by Pfanzagl in deriving his version of subjective probability. They are included here due to the difficulty in finding his book in print. The axioms and proofs are in chapter "12. Events, Utility, and Subjective Probability" [7] (pp. 195-220). No quotations are used and [any comments from the author will be bracketed].

12.1 **The Algebra of Events**

Let $\mathcal{E}$ be a system of events. These events will be denoted by capital letters like $P,\ Q,\ R,\ \cdots$. We will assume that for events in $\mathcal{E}$ the following three operations are defined: the *join* ($\cup$), the *meet* ($\cap$) and the *complement* (*negation*) $(\bar{\ })$. $P \cup Q$ is to be interpreted as the event which obtains if at least one of the events $P,\ Q$ obtains. $P \cap Q$ is to be interpreted as the event which obtains if both P and Q, obtain. $\bar{P}$ is to be interpreted as the event which obtains if P does not obtain. We shall assume that $\mathcal{E}$ is closed under these three options.

Furthermore, we shall assume that the following axioms are fulfilled:

12.1.1 **Definition:** A system $\mathcal{E}$ is a *Boolean algebra* if operations $\cup,\ \cap,\ \bar{\ }$ with the following properties are defined:

12.1.2 commutativity $\quad P \cap Q = Q \cap P, P \cup Q = Q \cup P$

12.1.3 associativity $\quad P \cup (Q \cup R) = (P \cup Q) \cup R$

$P \cap (Q \cap R) = (P \cap Q) \cap R$

12.1.4 distributivity $P \cap (Q \cup R) = (P \cap Q) \cup (Q \cap R)$
$P \cup (Q \cap R) = (P \cup Q) \cap (Q \cup R)$
12.1.5 absorption law $(P \cup Q) \cap Q = Q, (P \cap Q) \cup Q = Q$
12.1.6 $(P \cap \bar{P}) \cup Q = Q, (P \cup \bar{P}) \cap Q$=Q

This system of axioms remains unchanged if we interchange $\cap$ and $\cup$. Therefore, from the proof of any consequence of this system of axioms we obtain a proof of the "dual" consequence by interchanging $\cap$ and $\cup$. In the following we will prove only one of the two "dual" consequences.

In sections following 12.2, we will use the shorter expression PQ instead of $P \cap Q$.

…

Now we can define a binary relation between events by
12.1.8 **Definition:** $P \subset Q$ iff $P \cap Q = P$

By the absorption law (12.1.5), "$P \subset Q$ iff $P \cup Q = P$" is an equivalent definition.

The intuitive interpretation of this relation is that P implies Q, i.e. Q obtains if P obtains.

…

12.2 **The Space of Wagers**

…

The set of events will be denoted by $\mathcal{E}$, the set of possible outcomes by A. By a *simple wager* aPb with $a, b \in A$ and $P \in \mathcal{E}$, we mean the wager leading to outcome a if P obtains and to outcome b if P does not obtain.

For the sake of brevity let
$APA \coloneqq \{aPb: a, b \in A\}$ and $A\mathcal{E}A \coloneqq \{aPb: a, b \in A, P \in \mathcal{E}\}$.

About $\mathcal{E}$ and A we will make the following basic assumptions
12.2.1 $\mathcal{E}$ is a Boolean algebra.
12.2.2 A is an ordered and connected set containing at least two elements.

In the simplest case $\mathcal{E}$ might consist of the events $O, E, P, \overline{P}$, and A of different quantiles of a simple commodity (e.g. money) [measure], filling an interval.

Of basic importance for the following is
12.2.3 **Order Axiom:** $A\mathcal{E}A$ is an order system (3.2.1).

["3.2 **Order Relations**
3.2.1 **Definitions:** A relational system $\langle A, \approx, \prec \rangle$ is called an *order system* iff the following axioms are fulfilled:
3.2.2 **Order axioms:**

O1. For all $a, b \in A$, exactly one of the relations
$a \approx b, a \prec b, b \prec a$ holds
O2. "$\approx$" is an equivalence relation
O3. "$\prec$" is transitive.

A trivial example of an order system is $\langle \mathbb{R}, =, < \rangle$, the system of real number. Most empirical relational systems are order systems. If, for example, A, are tones, "$\approx$" and "$\prec$" might be equivalence and order according to pitch.

As in section 1.4, we can consider $\bar{\bar{A}}$, the set of all equivalence classes and the relational system $\langle \bar{\bar{A}}, =, < \rangle$ defined as follows:
$\bar{\bar{a}} = \bar{\bar{b}}$, if the two classes consist of the same elements.
$\bar{\bar{a}} < \bar{\bar{b}}$, if $a \prec b$."]

According to the order axiom, between any two elements $A\mathcal{E}A$ exactly one of the relations $\succ$, $\approx$ or $\prec$ holds. The intuitive meaning of this order according to utility [plausibility]. The relation $\approx$ is to be interpreted as an equivalence (in the sense of equal utility [plausibility]) not as an equality. Such an order relation may, for instance, be defined by means of the (objective) probability with which one element of $A\mathcal{E}A$ is [more plausible than] another, equivalence (preference [more plausible]) being the case in which this probability equals (is greater than) ½.

The order of $A\mathcal{E}A$ induces an order in $\mathcal{E}$, if the following axiom is fulfilled (see Savage (1954) [6], p. 31, p.4).

12.2.4 **Uniqueness Axiom:** If for a special pair $a_0, b_0 \in A$ with $a_0 > b_0$: $a_0 P b_0 \overset{\succ}{\underset{\prec}{\approx}} a_0 Q b_0$, then $aPb \overset{\succ}{\underset{\prec}{\approx}} aQb$ for all $a, b \in A$ with $a > b$.

If 12.2.4 is fulfilled, we may define an order in $\mathcal{E}$ by

12.2.5 **Definition:** $P \overset{\succ}{\underset{\prec}{\approx}} Q$ iff $aPb \overset{\succ}{\underset{\prec}{\approx}} aQb$ for all $a, b \in A$ with $a > b$.

We remark that under suitable assumptions the order defined in 12.2.5 refines the order defined by 12.1.8 (see 12.3.4).

A special role is played by the events equivalent to O, called "*almost impossible*", and the events equivalent to E, called "*almost sure*".

Interpreting $\prec$ on $A\mathcal{E}A$ as an order of wagers according to utility [plausibility] and $<$ on A as an order of outcomes according to utility [plausibility] we obtain as a formalized consequence of the principles a) and b) the following

12.2.6 **Postulate:**

a) $aPb \approx b\bar{P}a$ for all $P \in \mathcal{E}$ and $a, b \in A$,

b) $P \approx O$ implies $aPb \approx a'Pb$ for all $a, a', b \in A$ and
$P \not\approx O$ implies $aPb \prec a'Pb$ for all $a, a', b \in A$ with $a < a'$.

We remark that part (i) of the sure-thing principle [desiderata 3.c] is contained in 12.2.2: By assuming $<$ to be an order relation for which the equivalence relation is the identity, we identify outcomes of equal utility [plausibility].

As a consequence of 12.2.6 we obtain that

12.2.6.c) $P \approx O$ iff $\bar{P} \approx E$.

Proof: 12.2.5 and 12.2.6.a) imply that $\bar{P} \approx E$ iff $bPa \approx bOa$ for all $b < a$. By 12.2.6.b) this is equivalent to $P \approx O$.

12.2.7 **Continuity Axiom:**

$a \to aPb$ is continuous for all $b \in A$.

Together with 12.2.6.a) this implies that

$b \to aPb$ is continuous for all $a \in A$.

12.2.8 **Proposition:** The map of $A \times A$ into APA defined by $(a, b) \to aPb$ is continuous.

Proof: If P is almost impossible or almost sure, the assertion is trivial. If $P \not\approx O, E$ we argue as follows. As each one of the maps $a \to aPb$ and $b \to aPb$ is monotone increasing and continuous, $(a, b) \to aPb$ is continuous by 3.7.11.

A special role is played by the wagers aPa leading to outcome a regardless of whether P obtains or not. Though it seems most natural to require such wagers to be equivalent to a (for all $P \in \mathcal{E}$), it turns out to be unnecessary to formalize such an assumption in this section.

12.2.9 **Proposition:** The map $a \to aPa$ of A into APA is monotone increasing and continuous.

Proof: If P is almost impossible or almost sure the assertion is trivial. If $P \not\approx O, E$ we argue as follows: from 12.2.6 we obtain for $a < b$: $aPa \prec aPb \prec bPb$. Continuity follows immediately from the fact that $a \to (a, a)$ is continuous by (3.7.2) and $(a, a) \to aPa$ is continuous by 12.2.8.

12.2.10 **Theorem:** For all $a, b \in A$ with $a \neq b$ there exists exactly one c between a and b such that $cPc \approx aPb$.

Proof: Follows immediately from (12.2.9) and the intermediate value theorem (3.6.13).

12.2.11 **Definition:** "$\circ_P$" is the operation which assigns to each pair $a, b \in A$ an element $a \circ_P b \in A$ defined by $aPb \approx (a \circ_P b)P(a \circ_P b)$.

We remark that the existence of such an element is guaranteed by 12.2.10. $a \circ_P b$ will be called the *safety equivalent* of the wager aPb, for the subject is indifferent between the wager aPb and the amount of $a \circ_P b$.

12.2.12 **Theorem:** The operation $\circ_P$ and the relation $\approx$ have the following properties

(i) $\circ_P$ is intern except $P \approx O$ or $P \approx E$.

(ii) $a \circ_E b = a$.

(iii) $\circ_P$ is increasing in both variables except $P \approx O$ or $P \approx E$.

(iv) $(a, b) \to a \circ_P b$ is continuous.
(v) $a \circ_P b = b \circ_{\bar{P}} a$ for all $a, b \in A, P \in \mathcal{E}$.
(vi) The following three sentences are equivalent:
a) $a \circ_P b = a \circ_Q b$ for at least one pair $a, b \in A, a \neq b$,
b) $a \circ_P b = a \circ_Q b$ for all $a, b \in A$, and
c) $P \approx Q$.

Proof:

(i) See 12.2.10.
(ii) Follows immediately from $aEb \approx aEa$ (12.2.6).
(iii) We shall prove monotony in the first variable only: if $a' < a''$, 12.2.6 implies $a'Pb \prec a''Pb$. Let $c' = a' \circ_P b$, $c'' = a'' \circ_P b$. We have $c'Pc' \approx a'Pb \prec a''Pb \approx c''Pc''$. Hence 12.2.9 implies $c' < c''$.
(iv) We have to show that: $\{(a, b) \in A \times A: a \circ_P b < c\}$ is open for all $c \in A$. We have $a \circ_P b < c$ iff $aPb \approx (a \circ_P b)P(a \circ_P b) \prec cPc$. Hence $\{(a, b) \in A \times A: a \circ_P b < c\} = \{(a, b) \in A \times A: aPb \prec cPc\}$. As $(a, b) \to aPb$ is continuous (12.2.8), this is an open set for all $c \in A$. The other cases are dealt with similarly.
(v) Follows immediately from 12.2.6a).
(vi) Follows immediately from the uniqueness axiom (12.2.4) and (v).

12.3 **Compound Wagers**

…

12.3.2 **Postulate:** For all $a, b, c, d \in A$ and all $P, Q \in \mathcal{E}, (\not\approx O, E)$:

$\left(a \circ_{Q|P} b\right) \circ_P \left(c \circ_{Q|\bar{P}} d\right) = \left(a \circ_{P|Q} c\right) \circ_Q \left(b \circ_{P|\bar{Q}} d\right)$.

Furthermore the wagers $(a(Q|P)b)Pb$ and $aPQb$ are identical, as for both wagers $PQ \to a, P\bar{Q} \to b, \bar{P}Q \to b, \bar{P}\bar{Q} \to b$. Hence they are judged equivalent due to our lack of illusion principle [desiderata 3.a]. Furthermore we obtain from the sure thing principle [desiderata 2 and 3.c] $(a(Q|P)b)Pb \approx \left(a \circ_{Q|P} b\right)Pb \approx \left(a \circ_{Q|P} b\right) \circ_P b$ and $aPQb \approx a \circ_{PQ} b$. This suggests the following

12.3.3 **Postulate:** For all $a, b \in A$ and all $P, Q \in \mathcal{E}$ $(P \not\approx O)$:

$\left(a \circ_{Q|P} b\right) \circ_P b = a \circ_{PQ} b$.

12.3.4 **Theorem:** The order in $\mathcal{E}$ defined by 12.2.5 refines the order defined by 12.1.8: $P \supset Q$ implies $aPb \underset{\approx}{\succ} aQb$ for all $a, b \in A$ with $a > b$.

Proof: By definition 12.1.8, $P \supset Q$ is equivalent to $P \cap Q = Q$. Hence 12.3.3 implies $\left(a \circ_{Q|P} b\right) \circ_P b = a \circ_Q b$.

By 12.2.12 (ii) and (iii) $a > b$ implies $a \geq a \circ_{Q|P} b$. Hence $a \circ_P b \geq a \circ_Q b$ which proves the assertion.

In the following sections we shall only use properties of the operation $\circ_P$ which are stated in 12.2.12, 12.3.2 and 12.3.3. Instead of deriving these postulates from more general axioms (order, uniqueness and continuity) and principles (sure-thing and lack of illusion), another possible approach would be to forget about these general axioms and principles and to state 12.2.12, 12.3.2 and 12.3.3 together with the continuity axiom as fundamental axioms.

…

12.5 **Theorems on Utility and Subjective Probability**

12.5.1 **Definition:** The event $Q \in \mathcal{E}$ is *independent* of the event $P \in \mathcal{E}'$ iff $\circ_{Q|P} = \circ_{P|Q}$.

This definition is justified by the fact that $\circ_{Q|P} = \circ_{Q|\bar{P}}$ means that the knowledge of whether P or $\bar{P}$ obtains is irrelevant for the evaluation of wagers based on Q. It will be shown in 12.5.14 that independence is a symmetric property, i.e. if Q is independent of P, then also P is independent of Q.

12.5.2 **Axiom:** For each event $P \in \mathcal{E}'$ there exists an element of $Q \in \mathcal{E}'$ which is independent of P.

This axiom is not very strong, of course, because any rational person will be willing to consider the tossing of a coin as an event which is independent of all other uncertain events.

…

References

1. Ridley, M., *How Innovation Works: And Why It Flourishes in Freedom*. 1st ed. 2020, New York: Harper Collins Publishers.
2. Postrel, V., *The Fabric of Civilization: How Textiles Made the World*. 1st ed. 2020, New York: Basic Books.
3. Stanley, E.H., *Interview with Eugene H. Stanley*, in *IIM Kozhikode Society & Management Review*, K. Gangopadhyay, Editor. 2013, Sage Publication. p. 73-78.
4. Gibbs, J.W., *Elementary Principles in Statistical Mechanics Developed with Especial Reference to the Rational Foundation of Thermodynamics*. 1902, New York: Charles Scribner's Sons. 207.
5. Shannon, C.E., *A Mathematical Theory of Communication.* The Bell System Technical Journal, 1948. **27**: p. 379-423, 623-656.
6. Savage, L.J., *The Foundations of Statistics*. 2nd ed. 1972, New York: Dover Publications.
7. Pfanzagl, J., *Theory of Measurement*. 1968, New York: John Wiley and Sons.
8. Ellsberg, D., *Risk, Ambiguity, and the Savage Axioms.* The Quarterly Journal of Economics, 1961. **75**(4): p. 28.
9. Allais, P.M., *Le Comportement de l'Homme Rationnel Devant Le Risque- Critique des Postulats et Axiomes de L'Ecole Americaine.* Econometrica, 1953. **21**(4): p. 45.
10. Mises, L.v., *Human Action: A Treatise on Economics*. 1998, Auburn: Ludwig von Mises Institute.
11. Jaynes, E.T., *Information Theory and Statistical Mechanics.* The Physical Review, 1957. **106**(4): p. 620-630.
12. Matsoukas, T., *Generalized Statistical Thermodynamics*. 2018: Springer.
13. Rawls, J., *A Theory of Justice*. 2009: Harvard University Press.
14. Neumann, J.v. and O. Morgenstern, *Theory of Games and Economic Behavior*. 3rd ed. 1944: Princeton University Press.
15. Jaynes, E.T., *Probability Theory: The Logic of Science*. 1st ed. 2003: Cambridge University Press.
16. Jaynes, E.T., *Information Theory and Statistical Mechanics*. Brandeis University Summer Institute Lectures in Theoretical Physics, ed. K.W. Ford. Vol. 3 Statistical Physics. 1962, New York: W. A. Benjamin, Inc.
17. Smith, A., *The Wealth of Nations: An inquiry into the nature and causes of the Wealth of Nations*. 2010: Harriman House Limited.
18. Callen, H.B., *Thermodynamics and an Introduction to Thermostatistics*. 2nd ed. 1985, New York: John Wiley & Sons.
19. Abel, C. *Quantifying the Value of Bitcoin*. 2014; Available from: https://statisticaleconomics.org/2014/03/13/quantifying-the-value-of-bitcoin-2/.
20. Landau, L.D. and E.M. Lifshitz, *Statistical Physics, Parts 1 and 2*, in *Course of Theoretical Physics*. 1980, Butterworth-Heinemann, Oxford.
21. Ayers, R.U. and B. Warr, *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. 2009, Northhampton, MA: International Institute for Applied Systems Analysis.
22. *Annual Energy Review 2011*. 2012, Energy Information Agency.
23. *Gross Domestic Product: Implicit Price Deflator*. 2020 29 September 2020]; Available from: https://fred.stlouisfed.org/series/GDPDEF/.
24. *Consumer Price Index (CPI) Databases*. 2020 29 September 2020]; Available from: https://www.bls.gov/cpi/data.htm.
25. Yang, J. and W. Qiu, *Normalized Expected Utility-Entropy Measure of Risk.* Entropy, 2014. **16**.
26. Eddington, A.S., *The Nature of the Physical World.* Gifford Lectures, 1927. **74**.