INFORMATION RECOVERY IN COMPLEX ECONOMIC BEHAVIOR SYSTEMS

George Judge[1]

1 August, 2023

Abstract

In this paper, we recognize that reigning reductionist econometric paradigms do not provide a basis for understanding complex economic behavioral processes and systems. To acknowledge that the collective behavior in complex adaptive systems is different from that of the individual parts, we borrow some of the key fundamental concepts of entropy methods, and nonequilibrium systems dynamics to develop a theoretical economic framework for self-organizing-optimizing economic behavior systems at the macro-level that is based on nonlinear individual behavior at the micro-level. We illustrate the use of the proposed framework with applications that provide a basis for linking the econometric information sources.

Keywords. Adaptive Behavior; Emergence: Self-Organization; Causal entropy maximization; Information Theoretic Dynamics. Methods; Minimum Power Divergence; Statistical Equilibrium; Markov Processes.
JEL Classification: C1; C10; C2.

[1] George Judge, Professor of the Graduate School and Giannini Foundation, 207 Giannini Hall, University of California Berkeley, Berkeley, CA, 94720 (e-mail: gjudge@berkeley.edu). I would like to acknowledge the helpful interactions with H. Qian and the useful comments from M. Grendar, M. Henry, and S. Villas-Boas.

1. Introduction

In 1776 Adam Smith, in his book the *Wealth of Nations*, viewed the economic scene, and described the resulting emerging behavior system as being guided by an invisible hand based on self-interest. This self-organized concept, replaced the conventional reductionist concept, where the collected behavior of the whole is different, from that of its parts. In this paper we use the concepts of emergence and self-organization of what may seem like a seemingly random system, as a basis for information recovery in complex economic behavior systems that are stochastic, dynamic, nonlinear in nature, seldom if ever in equilibrium, These complex economic behavior systems have nonlinear causality, and the parts are interdependent. Although Boltzmann's equation (Boltzmann,1872) on the principles of statistical mechanics and the study of nonlinear systems occurred over a century ago, the study of complex economic behavior systems has not been a topic of interest in econometrics.

To acknowledge this void, we borrow some of the key concepts of entropy methods, quantum economics, and nonequilibrium emergent systems dynamics, to develop a theoretical economic framework for analyzing outcomes at the macro-level, based on self-organizing-optimizing economic behavior at the micro-level. In this context, emergence refers to the collective behavior in complex adaptive systems, that is not present in their individual parts. For relevant information on entropy and quantum economics see (Jakimowicz , 2020), and on emergence, see Arrow, Anderson and Pines,(1988), and Laughlin and Pines,(2002), and the references therein. In emergent self-organizing equilibrium seeking stochastic economic behavior systems, we indicate how to use information theoretic entropy-divergence methods to represent and evaluate the complex nature of the economic behavior systems, and to link economic-econometric information sources. Nonlinear probabilistic reasoning provides a natural language and serves as the scientific logic. A macroeconomic system's emergent behavior and predictability may then be studied quite naturally by information theoretic entropy divergence measures that reflect a state of knowledge in an inferential context.

Most complex economic behavior system theories are expressed in terms of reductionist models, that seek to understand the behavior of the whole by reducing them to the interaction of their parts. Thus, the premise of reductionism is that a complex system can be broken down into constituent parts, which are studied independently, and then reassembled to understand the behavior of the system as a whole. The econometric dynamic, stochastic, general equilibrium reductionist counterpart usually appears in a functional form-linear mode and uses observed data consistent empirical sample moments-constraints, such as

$$\mathbf{h}\left(\mathbf{Y},\mathbf{X},\mathbf{Z};\boldsymbol{\beta}\right)=n^{-1}\left[\mathbf{Z}'\left(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}\right)\right]\xrightarrow{p}\mathbf{0}, \tag{1.1}$$

where $\mathbf{Y},\mathbf{X}$, and $\mathbf{Z}$ are respectively a $n\times1$, $n\times k$, $n\times m$ vector/matrix of dependent variables, explanatory variables and instruments, and $\boldsymbol{\beta}$ is an unknown and unobservable parameter vector. A solution to the stochastic recovery problem (1.1) is usually based on parametric and semiparametric likelihood functions and on a range of observed data sampling processes that involve a finite set of moments. The moments are usually based on indirect noisy observations that are in the effect domain while our interest is in the causal domain. Higher order moments are assumed to be well behaved, an assumption often violated in real world markets where long tail distributions prevail. Obtaining a meaningful probabilistic basis for a realistic dynamic economic behavior system from this reductionist mix of model, data, and econometric method requires a computational task that is neither feasible nor meaningful.

In two current papers on the state of macro econometric models, Stiglitz, (2018) and Hendry and Meullbauer, (2018) make clear that reigning paradigms involving reductionist dynamic stochastic general equilibrium (DSGE) models do not provide a basis for understanding complex economic behavioral processes and systems and predicting deep changes in the economy. Given the unsatisfactory nature of the reductionist econometric information recovery process, in this paper we borrow the key concepts of self-organizing-emergent-optimizing economic behavior in nonequilibrium statistical-econometric systems and information theoretic Cressie-Read (CR) entropy-based methods to develop a recovery framework that is based on nonlinear stochastic dynamics at the macro level and stochastic individual behavior at the micro-level.

1.1 Structure of the Paper

In the sections ahead, we provide an information theoretic entropy-divergence framework for combining the traditional econometric and information sources. In Section 2 we discuss the underlying connection between economic behavior systems-processes and entropy and how this linkage helps establish an entropy driven statistical framework for studying self-organizing-optimizing economic behavior systems. In Section 3 we discuss the information theoretic Cressie-Read (CR) divergence function framework for combining information sources-processes. In section 4 we demonstrate a simple economic information recovery problem that may provide a basis for linking the econometric learning information sources. In section 5 we consider a useful attribute of the CR divergence function, that leads to a Markov process combining framework in discrete state and time and the Kullback-Leibler basis for expressing the mutual information of two processes In section 6 we summarize the paper and look ahead and speculate on possibilities for measuring information transfer and interaction in systems.

2. Self-Organized Economic Behavior-Entropy Connection

In the attempt to formulate statistical physics, in a way that is independent of the physics of particles, the Maximum Entropy Principle (MEP) was developed by Jaynes (1978). In the MEP context, entropy is used as a means of statistical inference on multinomial processes. Thus MEP is closely related to the question of finding the most likely observable microstates of a system or process, where a microstate is a particular configuration of the components of a system. The distribution function p, or the histogram k, is such a macro state. The focus is on finding the most likely observable distribution function of a given process of a system.

Georgescu-Roegen in (1971 noted that a theoretical mathematical model of the economic behavior system, had something in common with random processes that obeyed the laws of probability-stochastic processes and entropy. Recently, Wissner-Gross and Freer, (2013) exhibited a connection between adaptive intelligent behavior and causal entropy maximization that provided an entropy basis for self-organization(emergence)-optimization in multi-agent behavioral economic systems. This causal entropic force connection, that is consistent with the idea that an economic system adapts behavior in line with an optimizing principle (see Judge, (2015), Miller and Judge, (2015) and Judge, (2016)), leads to self-organized equilibrium seeking behavior in an open seemingly random economic behavioral system and establishes an entropy driven inferential framework. In this context behavior-market systems are equilibrium stationary state seeking and the entropic force quantifies the intuitive notion that systems are more likely to move from a low probability to a high probability state. Predictability of a system in terms of an entropy function, or entropy functions, is then equivalent to the study of its statistical-nonequilibrium nature, Thus, for example, if we characterize a macro state by its probability distribution-multiplicity of microstates, entropy reflects the number of ways a macro state can evolve along a path of possible microstates. This view permits us to recognize that economic data comes from systems with dynamic adaptive behavior that are non-deterministic in nature, involve uncertainties that can be quantified by the notion of information and are driven toward a certain stationary state associated with a functional and hierarchical structure. As we seek new ways to think about the causal adaptive behavior of large complex and dynamic economic systems, the Wissner-Gross and Freer (2013) economic behavior entropy connection means that we can use entropy as the systems status measure-optimizing criterion. A uniform-unstructured distribution of the microstates corresponds to a macro state with maximum entropy and minimum system information. In this context, causal entropy maximization is a link that leads us to believe that an economic-behavioral system with a large number of agents interacting locally and in finite time is in fact optimizing itself. One of the most general entropy information-theoretical functions that measure uncertainties, or missing information, or discrimination, is a CR family of divergence-entropy measures and this is the topic to which we turn.

3. The Information Theoretic Minimum Divergence Family

In this section, we discuss how information theoretic entropy methods provide a natural basis in questions regarding economic behavioral systems, for establishing a causal influence-econometric-inferential link to the data and solving the resulting statistical equilibrium problem. In this information recovery context, a natural solution in terms of economic policy is to use estimation and inference methods that are designed to deal with systems that may be nonlinear and stochastic in nature, where uncertainty and random behavior are basic to information recovery. To identify estimation and inference measures that represent a way to link the model of the process to a family of possible likelihood functions, we use the Cressie and Read (CR), (1984) and Read and Cressie, (1988) single parameter CR family of entropic function-power divergence measures given by

$$I\left(\mathbf{p},\mathbf{q},\gamma\right)=\frac{1}{\gamma\left(\gamma+1\right)}\sum_{i=1}^{n}p_i\left[\left(\frac{p_i}{q_i}\right)^{\gamma}-1\right]. \qquad (3.1)$$

In (3.1), the value of γ is a parameter that indexes members of the CR family, the p_i's represent the subject data probability distribution, and the q_i's are reference non sample prior probabilities and **p** and **q** are $n\times1$ vectors of p_i's and q_i's, respectively. The usual probability distribution characteristics of $p_i, q_i \in [0,1]\ \forall i$, $\sum_{i=1}^{n} p_i = 1$, and $\sum_{i=1}^{n} q_i = 1$ are assumed to hold. In (3.1), as γ varies, the resulting CR-entropy statistical family of estimators that minimize power divergence, exhibit qualitatively different sampling behavior that includes maximum entropy and in general, a range of additive and correlated systems (see Gorban, et al. (2010), Judge and Mittelhammer (2011), (2012)). In identifying the probability space, the CR family of power divergences is defined through a class of additive convex functions that represents *a* *broad family of likelihood functional relationships* and test statistics. All well-known divergences-entropies belong to this class of CR functions. In addition, the CR measure exhibits formal convexity in **p**, for all values of γ and **q**, and embodies the required probability system characteristics, such as invariance with respect to a monotonic transformation of the divergence measures. As γ varies, the optimized value of, $I(\mathbf{p},\mathbf{q},\gamma)$ represents a range of data sampling processes, likelihood functions and ensembles, and the corresponding estimators that minimize power divergence exhibit qualitatively different sampling behavior. For example, in the limit as $\gamma \to 0$, the solution of the first-order condition leads to the maximum entropy and the logistic expression for the conditional probabilities. In the limit $\gamma \to 0$, the Maximum Exponential Empirical Likelihood (MEEL) estimator,is the most likely distribution to be observed from a statistical or combinatorial point of view

4. Entropy Based Probability Density Functions

To develop an economic framework for self-organizing-optimizing economic behavior systems at the macro-level, that is based on individual behavior at the micro-level that describes the microstate in the economic system, we can use the CR- $I(\mathbf{p},\mathbf{q},\gamma)$, $\gamma \to 0$ divergence-distance measure and a sample of micro data. As an example, in this section, assuming no prior information, we use a sample of a country's income data and discretize income levels by a finite number of nonoverlapping intervals-histograms, to establish a data-based link to recover the micro income probability density function-distribution and the macro income inequality measure. We use histograms to connect the micro and macro states, The microstates that are the subject of inference and the number of microstates-micro configurations, is called the multiplicity of the macro state. In recovering the income probability density function-distribution from a sample of N positive real numbers, we assume the income probability to be represented by K histograms that span the income sample space. Samples from these partitions yield histogram outcomes of the discrete random income variable d_j, for $j = 1, 2, \ldots, n$, and under repeated observation, one of n histograms-micro configurations associated with the macro state income is observed with probability p_j. After a large number of trials, we recover first-moment sample information in the form of the mean value of $\sum_{j=1}^{K} d_j\, p_j = \bar{d}$. Under this specification, when in the limit γ→0, the CR $I(\boldsymbol{p}, \boldsymbol{q}, \gamma)$ converges to an estimation criterion equivalent to the maximum exponential empirical likelihood (MEEL) metric $H(\boldsymbol{p}) = -\sum_{j=1}^{K} p_j \ln(p_j)$. The extremum problem likelihood-entropy function may then be formulated as

$$\max_{\boldsymbol{p}} \left[-\sum_{j=1}^{K} p_j \ln\, p_j \;\middle|\; \sum_{j=1}^{K} p_j\, d_j = \bar{d}, \sum_{j=1}^{K} p_j = 1, p_j > 0, \text{ for all j} \right]. \tag{4.1}$$

The corresponding Lagrange function-extremum problem is

$$L(\boldsymbol{p}, \eta, \boldsymbol{\lambda}) \equiv -\sum_{j=1}^{K} p_j \ln(p_j) + \boldsymbol{\lambda}\left(\bar{d} - \sum_{j=1}^{K} p_j d_j\right) + \eta\left(1 - \sum_{j=1}^{K} p_j\right). \tag{4.2}$$

Solving the first-order conditions yields the exponential result

$$\hat{p}_j = \frac{exp\,(-d_j\hat{\boldsymbol{\lambda}})}{\sum_{j=1}^{K} exp\,(-d_j\hat{\boldsymbol{\lambda}})} \tag{4.3}$$

for the jth income outcome and the mean-related income distribution. Using the CR ($\gamma \to 0$) entropy functional, the joint histograms and the mean of a country's income data, the resulting probability density income distribution and the entropy inequality income measure $\sum_{j=1}^{K} p_j \ln(p_j)$ *is* recovered, and provides a measure of income inequality. For an application of this income distribution information recovery method, see Villas Boas, et al., (2019).

4.1 **Reference Distribution**

The income probability density function-distribution reported in section 4.1, was obtained using a uniform reference distributions **q.** In practice when estimating income distributions, in many situations we may also have information from other countries, which we denote as income patterns and probabilities $\mathbf{p} = (\mathrm{p}_1, \mathrm{p}_2, \cdots, \mathrm{p}_\mathrm{k})'$, in the form of a reference distribution of probabilities $\mathbf{q} = (\mathrm{q}_1, \mathrm{q}_2, \cdots, \mathrm{q}_\mathrm{k})'$. That is, when making use of the sample of micro data that enters in the consistency relations, in (4.1) and (4.2), there may be additional information in the form of a non-uniform reference distribution $\boldsymbol{q}$, that provides information concerning the income probability distribution. When such prior non-uniform reference information-knowledge exists, we may wish to follow Kullback and Liebler (1951), Kullback, (1959) and Good, (1963) and incorporate this information in the form of the principle of minimum cross-relative entropy or Kullback-Liebler (KL) directed divergence. This minimal discriminability principle implies one would choose, given the constraints, the estimate of **p** that can be discriminated from the non-uniform reference distribution **q**, with a minimum of difference. The KL estimation objective function can be seen as a particular case of the more general CR family of entropy discrepancy-distance measures introduced in section 3. Thus, the principle of minimum cross-relative entropy or KL directed divergence, makes it possible to use-combine the information to recover a country's income distributions and entropy measure of inequality. The divergence functions over probability distributions, provide a strong and rigorous macro-economic index for systems of various behaviors.

For example, instead of a uniform distribution **q,** we can make use of the distribution q_{ML}, as the reference distribution in (3.1). In the CR formulation (3.1), $\gamma \to -1$ and probabilities q_{ML} replaces the uniform reference distribution in (3.1). This leads to the KL empirical likelihood criterion

$$\lim_{\gamma \to -1} I(\boldsymbol{p}, \boldsymbol{q_{ML}}, \gamma) = \sum_{j=1}^{n} q_{jML} \ln(p_j / q_{jML}) = \sum_{j=1}^{n} q_{jML} \ln(p_j) - \sum_{j=1}^{n} q_{jML} \ln\left(q_{jML}\right),$$

(4.4)

where $\sum_{j=1}^{n} q_{jML} \ln\left(q_{jML}\right)$, is a constant. Using this criterion, the data constraint, and a selected mean, results in

$$\hat{p}_{jML}\left(\bar{d}, \hat{\lambda}\right) = q_{jML} \left(1 + \hat{\lambda}\left(d_j - \bar{d}\right)\right)^{-1} \qquad \text{for } j = 1, \dots, n, \qquad (4.5)$$

where $\hat{\lambda}$ is such that $\hat{p}_{jML}\left(\bar{d}, \hat{\lambda}\right)$ satisfies the mean constraint (5.1). The solution, $\hat{p}_{jML}$, may then be used to estimate the combined income distribution and a new entropy measure of income inequality. For empirical applications of this type of information recovery from micro data, and the inequality income measure, see Villa-Boas, et al. (2019), and Judge (2019).

5. An Entropic Prior

Given this base, we now discuss another useful attribute of the divergence functions of the distributions. Assume we are observing an economic system on a relatively long, time scale. Then the current state of the system depends largely on the closest previous state, and the history over longer terms does not have much influence. After discretizing income levels, a Markov process with finite discrete state and discrete time is obtained. We denote this Markov process with the Markov matrix P. We can then demonstrate a monotonic decreasing of the CR entropy family (3.1), $\gamma \geq 0$, with Markov dynamics, where p(t) and q(t) are two distributions following the Markov chain, with different initial distribution p(0) and q(0):

$$\frac{p_i(t+1)}{q_i(t+1)} = \frac{\sum_{j=1} p_j(t)P_{ji}}{q_i(t+1)} = \frac{1}{q_i(t+1)} \sum_{\{j=1\}}^{n} \frac{p_j(t) q_j(t) P_{ji}}{q_j(t)}$$

$$= \frac{1}{\sum_{j=1}^{n} q_j(t)P_{ji}} \sum_{j=1}^{n} q_j(t)P_{ji} \left(\frac{p_j(t)}{q_j(t)}\right).$$

Then with a convex function $\varphi(x)$:

$$q_i(t+1)\varphi\left(\frac{p_i(t+1)}{q_i(t+1)}\right) \leq \frac{q_i(t+1)}{\sum_{j=1}^{n} q_j(t)P_{ji}} \sum_{j}^{n} q_{-j}(t)P_{ji}\varphi\left(\frac{p_j(t)}{q_j(t)}\right) = \sum_{j=1}^{n} q_j(t)P_{ji}\varphi\left(\frac{p_j(t)}{q_j(t)}\right).$$

The function $\varphi(x) = x^{\gamma+1}$ is convex and additive. Therefore,

$$I(\{p_i(t+1)\},\{q_i(t+1)\},\gamma) = \frac{1}{\gamma}\left\{\sum_{i=1}^{n} q_i(t+1)\left(\frac{p_i(t+1)}{q_i(t+1)}\right)^{\gamma+1}\right\}$$

$$\leq \frac{1}{\gamma}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} q_j(t)\, P_{ji}\left(\frac{p_j(t)}{q_j(t)}\right)^{\gamma+1} - 1\right\}$$

$$= \frac{1}{\gamma}\left\{\sum_{j=1}^{n} q_j(t)\left(\frac{p_j(t)}{q_j(t)}\right)^{\gamma+1} - 1\right\} = I(\{p_i(t)\},\{q_i(t)\},\gamma), \qquad (5.1)$$

where the inequality in (5.1) follows from the convexity of $(\cdot)^{\gamma+1}$.

In many cases, we are only interested in a periodic irreducible Markov processes, because the probability of a state being strictly periodic is practically zero, and if the entire system can be reduced into subsystems, we can study the subsystems individually. In those cases, the CR entropy approaches zero as *p*, *q* approaches the unique stationary distribution. Finally, the limit of *I(**p**, **q**, γ)* when γ goes to 0 is

$$\lim_{\gamma \to 0} I(\{p_i\}, \{q_i\}, \gamma) = \sum_{i=1}^{n} ln\left(\frac{p_i}{q_i}\right), \qquad (5.2)$$

which is the KL divergence measure in the discrete case. This criterion leads to a natural measure of the deviation of the distribution of probabilities **p** and **q**. Under the principle of minimum discriminability, the difference $I(\boldsymbol{p}, \boldsymbol{q})$ is minimized. To take account of both an informative prior reference distribution, and the micro data sample information, the minimum cross-entropy solution may be obtained from the minimization problem (5.1), subject to the moment consistency constraints and the adding up-normalization constraint. The entropy measure gives a monotonicity relation to the Markov process., and in contrast to moment-based estimation problems, where matching the leading moments does not guarantee convergence in probability measures, KL-divergence upper bounds the large deviation rate function, and thus controls moments of all orders. Consequently, the $I(\mathbf{p}, \mathbf{q}, \gamma)$ or $I(\mathbf{q}, \mathbf{p}, \gamma)$ divergence functions over discrete probability distributions defined on the same probability space, provide a strong and rigorous macro-economic index for systems with various behaviors. Importantly **in** terms of information recovery, this entropy measure provides a divergence framework, and a strong and rigorous basis for integrating-linking the transfer of information, for discrete probability distributions defined on the same probability space, and that come from various sources-behaviors.

6. Summing Up

This paper contains examples of the use of entropy in recovering information in complex economic behavior systems. We recognize the failure of reigning paradigms to provide a basis for understanding the hidden nonlinear dynamics generated from complex economic behavioral processes and systems. To fill this void, we borrow some of the key concepts of entropy based nonlinear divergence methods and physical systems dynamics to suggest a framework for self-organizing-optimizing macroeconomic behavior systems that is based on nonlinear stochastic dynamics of individual behavior at the micro-level. A Markov decision process and a transitional reference distribution are used to measure the impact of the use of prior information on related macro-outcome variables. An entropy based minimum divergence method is used in a stochastic Markov dynamic context as a basis for predicting macroeconomic behavior.

Looking ahead, we recognize that for such things as climate policy (Kupers, (2020)), it is important to develop other nonlinear econometric frameworks for self-organizing-optimizing economic behavior systems at the macro-level that are based on nonlinear stochastic dynamics of individual behavior at the micro-level. Causality and the classification of policy changes and the stability of hidden structures within a paradigm are subjects to be addressed in future work.

References

Anderson, P.W., Arrow, K., and D. Pines, 1988, *The Economy As An Evolving Complex System*, Addison-Wesley, Reading, Ma.

Arthur, W., 2021, “Foundations of complexity economics,” *Nature Review Physics*, 3:136-145.

Boltzmann, L. Weitere studien uber das w ̈arme gleichgenicht unfer gasmol ̈akuler. Sitzungsberichte der Akademie der Wissenschaften 66 (1872), 265–370. Translation: Further studies on the thermal equilibrium of gas molecules, in Kinetic Theory 2, 88174, Ed. S.G. Brush, Pergamon, Oxford (1966).

Cressie, N. and T. Read, 1984. Multinomial Goodness of Fit tests. Journal of The Royal Statistical Society. Series B. 46(3): 440-464.

Georgescu-Roegen, N. 1971. The Entropy Law and the Economic Process. Harvard University Press, Harvard.

Good, I., 1963. Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables. The Annals of Mathematical Statistics, 34(3): 911-934

Gorban, A., P. Gorban, A. and G. Judge, 2010. Entropy: The Markov Ordering Approach. Entropy 12(5),:1145-1193.

Hendry, D. and J. Meullbauer, 2018. The Future of Macroeconomics: Macro Theory and Models at the Bank of England, Oxford Review of Economic Policy, 34(1-2):287-328.

Jakimowicz, A., 2020. The R ole of Entropy in The development of Economics, Entropy, 22(4): 452

Jaynes, E.T., 1978, Where do we stand on maximum entropy? In the Maximum Entropy Formalism, Ed. R.D. Levine and M. Tribus, pp: 15 – 118, Cambridge, Mass: MIT Press.

Judge, G.G., R.C. Mittelhammer, 2011. An Information Theoretic Approach to Econometrics. Cambridge University Press.

Judge, G.G., and R.C. Mittelhammer, 2012. Implications of the Cressie-Read Family of Additive Divergences for Information Recovery. Entropy 14(12):2427-2438.

Judge, G., 2015 Entropy Maximization as a Basis for Information Recovery in Dynamic Economic Behavioral Systems, Econometrics, 3(1):91-100.

Judge, G., 2016, Econometric Information Recovery in Behavioral Networks, Econometrics, 4:38. doi: 10.3390/econometrics4030038

Judge, G, 2019, Micro-Macro Connected Stochastic Dynamic Economic Behavior Systems. Econometrics, 6, 46; doi:10.3390/econometrics6040046.

Kullback, S. and R. Liebler, 1951. On Information and Sufficiency, Annals of Math. Stat., 22:79-86.

Kullback, J., 1959. Information Theory and Statistics, John Wiley, New York.

Kupers, R., 2020, A Climate Policy Revolution, Harvard University Press, Cambridge.

Laughlin R. and Pines, D., (2002), The Theory of Everything. PNAS 97:28.

Miller, D. and G. Judge, 2015, Information Recovery in a Dynamic Statistical Markov Model, Econometrics, 3:187-198.

Read, T.R. and N.A., Cressie, 1988. Goodness of Fit Statistics for Discrete Multivariate Data, New York: Springer Verlag.

Stiglitz, J., 2018. Where Modern Macroeconomics Went Wrong, Oxford Review of Economic Policy, 34(1-2):70-106.

Villas-Boas, S., Q. Fu and G. Judge, 2019. Entropy Based European Income Distributions and Inequality Measures, Physica A, Volume 514: 686-696.

Wissner-Gross, A.D and C. E. Freer, 2013. Causal Entropic Forces, Physical Review Letters, 110(16):167802-1-5.