UC Berkeley
CUDARE Working Papers

Title
Some Thoughts on Econometric Information Recovery

Permalink
https://escholarship.org/uc/item/5h048626

Author
Judge, George G.

Publication Date
2013-08-21

eScholarship.org Powered by the California Digital Library
University of California

University of California, Berkeley
Department of Agricultural &
Resource Economics

CUDARE Working Papers

Year 2013 *Paper 1135*

Some Thoughts on Econometric Information Recovery

George G. Judge

Copyright © 2013 by author(s).

Some Thoughts on Econometric Information Recovery

George G. Judge,[1]* University of California, Berkeley

[1]George G. Judge, Professor of ARE and the Graduate School, 207 Giannini Hall, University of California, Berkeley, Berkeley, CA, 94720 (e-mail: gjudge@berkeley.edu). It is a pleasure to acknowledge the helpful comments of S. Berto Villas-Boas, W. Cho, M. George, J. Geweke, D. Hendry, C. Hill, K. Judd, J. Lee, R. Mittelhammer, H. Pesaran, P. Phillips, A. Plastino, L. Schechter, R. Shiller, J. Stiglitz, S. Turnovsky, and A. Zaman. Of course they are not to be held responsible for this set of words and other symbols.

An Information Theoretic Approach To Econometric Information Recovery

ABSTRACT

This paper is concerned with the problem of information recovery and measuring evidence that involves uncontrolled indirect noisy effects data and stochastic ill posed inverse problems in economics-econometrics. Information theoretic methods based on a multi parametric family of power divergence measures are suggested as an estimation and inference framework for dealing with these problems, analyzing questions of a causal nature and learning about hidden dynamic economic processes and systems that may or may not be in equilibrium. The paper concludes with some comments on the implications for information recovery of continuing to use traditional economic-econometric models and methods.

Key words: Information theoretic methods, First order Markov processes, Inverse problems, Dynamic economic systems, Information Recovery.

JEL Classification: C40, C51

8/21/13

1. INTRODUCTION

The outcome of economic-behavioral processes and systems involves the rich ingredients of complexity, uncertainty, volatility and ambiguity. Faced with these challenges, conventional econometric models and methods have told us more about the unpredictability of the interdependent nature of economic processes, than how to understand and predict them. The statistical complexity of information recovery emerges, because economic systems are dynamic and seldom in equilibrium and there is no unique time invariant econometric model. Given an uncontrolled sample of effects data that is indirect, noisy and observational in nature, the focus of this paper is on how to use this information and information theoretic methods to choose a member from the family of economic and econometric models that reflects the unknown initial condition and the hidden dynamics of the economic process.

The organization of the paper is as follows: Section 2 contains a discussion of why traditional econometric methods are not appropriate for answering many economic information recovery questions. In Sections 3 and 4, estimation and inference methods are suggested as a basis for dealing with the dynamic stochastic inverse problems often faced in econometrics. Finally in Section 5 we note the implications of continuing to use traditional econometric models and methods that are not relevant for the data and economic questions at hand.

2. THE INFORMATION RECOVERY PROBLEM

In attempts to learn from a sample of indirect noisy data, the traditional approach in econometrics usually has a "what is" focus—what is the model, what are the parameters, and what is the sampling distribution underlying the sample of data. This classical approach to information recovery is consistent with traditional semiparametric and parametric econometric approaches to estimation and inference, from a finite sample of data. There is something that is troubling about this approach as an econometric basis for learning from a sample of OBSERVED data. To ask what the parameter or underlying likelihood-density function IS, rather than what it could be, would seem in many cases to be asking an incorrect and unnecessarily difficult question. The underlying econometric model framework is conceptual in nature and seldom, if ever, correctly specified. The parameters, structural or otherwise, are unobserved and indeed are unobservable. The sample information is based on indirect noisy sample observations, that come from an economic systems or processes that, are in general, not in equilibrium. Thus with uncertainty regarding existing conditions, the observations may not come from a single sampling distribution, but many sampling distributions and a micro canonical ensemble.

Most conceptual micro and macro economic models are single valued equilibrium in nature. Unfortunately in practice, economic processes and systems are stochastic in nature. Although, there is only one sampling distribution consistent with an economic system in equilibrium, there are a large number of possible ways an economic process-system may be out of equilibrium. In this situation it seems inappropriate to ask what the model is, but to ask what it could be. For many econometric problems and data sets, the

natural solution may not be a fixed distribution, but a set of distributions, each with its own probability. In this situation the value of an econometric prediction is related to how far the system is out of equilibrium.

The resulting uncertainty about existing conditions and the dynamics of the process, create problems for econometric model specification and make it difficult, using traditional direct econometric methods, to capture the underlying hidden structure or key dynamic functions of the economic system. Although economic processes are, in general, simple in nature, the underlying dynamics are complicated and not well understood. The result is a family of economic models, each incorrectly specified and containing inadequacies that provide an imperfect link to the indirect noisy observational data. Given sufficient parameterization it is always possible to match a model to any data set, without capturing the underlying hidden structure or the key functionality of the system at hand.

The data associated with information recovery in a dynamic economic system often consists of a sample of indirect noisy effects data. This type of data makes it is impossible to distinguish between mutual influence and causal influence and does not contain dynamic or directional information. Even introducing a lag in the mutual observations fails to distinguish information that is actually exchanged, from shared information and does not support time causality.

The indirect noisy observations that are used in an attempt to identify the underlying dynamic system and to measure causal influence, requires the solution of a stochastic inverse problem that is the inverse of the conventional forward problem, which

relates the model parameters to the observed data. Thus the data are in the effects domain, and our interest lies in the causal domain. If the number of measurements-data points, are smaller than the number of unknown parameters to be estimated, the stochastic inverse problem is in addition ill-posed. Without a large number of assumptions, the resulting stochastic ill-posed underdetermined inverse problem cannot be solved by traditional estimation and inference methods. As a result, conventional semi and parametric estimation and inference methods are fragile under this type of model and data uncertainty and in general are not applicable for answering causal influence dynamic economic system questions. Application of traditional econometric models and methods that are not suited to the ill-posed inverse information recovery task, has led to what Caballero (2010) has called a pretense of knowledge syndrome.

3. POSSIBLE INFORMATION THEORETIC SOLUTIONS

Uncertainty regarding the makeup and underlying dynamics of the economic model and the unknown data sampling probability process create unsolved problems as they relate to econometric information recovery. A natural solution is to make use of estimation and inference methods that are designed to deal with the nature of economic models and data and the resulting stochastic ill-posed inverse problem. In this context the Cressie and Read (1984 (CR)) family of likelihoods-functionals, provide a basis for linking the data and the macro and micro model parameters. This permits the researcher to exploit the statistical machinery of information theory to gain insights relative to the underlying causal behavior of a dynamic process from a sample of data from a system that may not

be in equilibrium. In developing this information theoretic econometric approach to estimation and inference, the CR single parameter family of informational functionals, represents a way to link the family of possible likelihood functions associated with the underlying sample of data, to the sampling model of the process. Information functionals of this type have an intuitive interpretation that reflects uncertainty as it relates to a model of the process and a model of the data for possibly out of equilibrium processes. This gives new meaning to what is a likelihood function and what is the appropriate way to represent the possible underlying sampling distribution and relate it to the economic-model parameters.

One possibility for implementing this approach is to use estimating equations-moment conditions (prior information) to model the process and link to the data. Discrete members of the CR family would then be used to model the data and identify the weighting of the possible density-likelihood functions. The outcome reflects in a probability sense, what we know about the unknown parameters and a possible density function. The result may be a single distribution or a canonical hyper distribution density of possible underlying distributions. An advantage of this approach, in addition to its optimality base, is that it permits the possibility of non exponential-heavy tails and volatile distributions. Given the problem and possibility space outlined in the previous paragraphs, the importance of developing probabilistic and predictive results from a connected-mutual sample of indirect noisy effects data, guides the information theoretic estimation and inference framework sketched in the section to follow.

Many of the criticisms directed to traditional frequentist econometric methods, can also be directed to traditional Bayesian methods of estimation and inference. For a first step toward the development of information theoretic methods in a Bayesian context, see Grendar and Judge (2009).

4. COPING WITH THE UNCERTAINTY OF UNCERTAINTY

In identifying estimation and inference measures that may be used as a basis for characterizing the data sampling process for indirect noisy observed data outcomes, we begin with the CR multi parametric family of goodness of fit-power divergence measures:

$$I(\mathbf{p},\mathbf{q},\gamma)=\frac{1}{\gamma(\gamma+1)}\sum_{i=1}^{n}p_i\left[\left(\frac{p_i}{q_i}\right)^{\gamma}-1\right]. \tag{4.1}$$

In (4.1), γ is a parameter that indexes members of the CR family, p_i's represent the subject probabilities and the q_i's, are interpreted as reference probabilities. Being probabilities, 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 (4.1), as γ varies, the resulting CR family of estimators that minimize power divergence, exhibit qualitatively different sampling behavior.

4.1 Behavioral Econometric Models

In traditional subject matter econometric models if we use a moment representation that is often used in the application of the general Method of moments, one

might use empirical sample moments such as $h(Y, X, Z; \beta) = n^{-1}\left[Z'(Y - X\beta) \right] = 0,$ as constraints. In this case, which is but one of an infinite set of possible orthogonality conditions, a solution to the stochastic inverse problem, for any given choice of the γ parameter, may be formulated as the following extremum-type estimator for $\boldsymbol{\beta}$:

$$\hat{\boldsymbol{\beta}}(\gamma) = \arg\min_{\boldsymbol{\beta} \in \mathbf{B}} \left[\min_{\mathbf{p}} \left\{ I(\mathbf{p}, \mathbf{q}, \gamma) \mid \sum_{i=1}^{n} p_i \mathbf{z}'_{i.} \left(y_i - \mathbf{x}_{i.} \boldsymbol{\beta} \right) = \mathbf{0}, \sum_{i=1}^{n} p_i = 1, p_i \geq 0 \forall i \right\} \right], \quad (4.2)$$

where **q** is usually taken as a uniform distribution. This class of estimation procedures is referred to as *Minimum Power Divergence* (MPD) estimation (see Gorban, Gorban, and Judge 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 and the CR power divergence measure encompasses a family of test statistics and leads to *a broad family of likelihood functions* within a moments-based estimation context. In the context of extremum metrics, maximum likelihood is embedded in the general CR (1984) family of power divergence statistics. This family represents a flexible set of pseudo-distance measures from which to derive empirical probabilities associated with the indirect noisy micro and macro data. This class of estimation procedures is referred to as *Minimum Power Divergence* (MPD) estimation (see Gorban et al. 2010; Judge and Mittelhammer 2011, 2012). The CR family of divergence measures (4.1), permits us to exploit the statistical machinery of information theory to gain an insight into the probability density function (PDF) behavior of dynamic economic systems and processes. The likelihood functionals-PDFs-divergences have an intuitive interpretation in terms of uncertainty and measures of

distance. To make use of the CR family of divergence measures to choose an optimal probability system, one might focus on minimizing a loss function in choosing gamma and follow Gorban (1984) and Gorban and Karlin (2003) and consider a parametric family of convex information divergences, which satisfy additivity and trace conditions. For example, convex combinations of $\mathrm{CR}(\gamma \geq 0)$ and $\mathrm{CR}(\gamma \geq -1)$ span an important part of the probability space and produce a remarkable family of divergences-distributions. *This generalized family of divergence measures permits a broadening of the canonical distribution functions and provides a framework for developing a loss-minimizing estimation rule.*

4.2 A Stochastic State Space Framework

Information theoretic dynamic economic models appears naturally and can be given a directional meaning in a conditional Markov framework, when state spaces and transition probabilities are introduced in the CR-MPD framework (Miller and Judge 2012). Looking at the scene in terms of a stochastic process lets us get a peek at some of the hidden dynamics of the system and introduce the role of time and probabilistic causality. The condition of probabilistic causality introduces an arrow of time restriction, and the Markov process satisfies this restriction. For example, if the decision outcomes exhibit first-order Markov character, the dynamic behavior of the agents may be represented by conditional transition probabilities $p(j,k,t)$, which represent the probability that agent *i* moves from state *j*=1, 2, …, *K*, to state *k* at a time *t*. Given observations on the micro behavior *Y*(*i*,*k*,*t*), the conditional discrete Markov decision

process framework may be used to model the agent-specific dynamic economic behavior which varies with t. In the conditional case, we have $T\times(K-1)\times(K-1)$ unknown transition probabilities in the $(K-1)\times(T-1)$ estimating equations. Information recovery goes from the data to the parameters, thus this is an ill-posed inverse estimation problem and traditional estimation methods are not applicable.

Linking the sample analog of the Markov process to the indirect noisy observations

$$Y(k,t)=\sum_{j=1}^{K}Y(j,t-1)\,p(j,k,t)+e(k,t), \tag{4.3}$$

leads to a new class of conditional Markov models that is based on a set of estimating equations-moment equations

$$\mathrm{E}\left[z_t'e(k,t)\right]=0, \tag{4.4}$$

where $\mathbf{z}_t$ is an appropriate set of instrumental system or intervention variables. By substituting of (4.5) into (4.6), we form a set of estimating equations expressed in terms of the unknown transition probabilities. As a basis for identifying parametric data sampling distributions and likelihood functions in the form of distances in probability space we again use the CR family of power divergence measures

$$I(p,q,\gamma)=\frac{2}{\gamma(1+\gamma)}\sum_{t=1}^{T}\sum_{j=1}^{K}\sum_{k=1}^{K}p(j,k,t)\left[\left(\frac{p(j,k,t)}{q(j,k,t)}\right)^{\gamma}-1\right], \tag{4.5}$$

to provide access to a rich set of distribution functions that encompasses a family of estimation objective functions indexed by discrete probability distributions that are

convex in *p*. Formally, the MPD problem may be solved by choosing transition probabilities *p* to minimize $I(p,q,\gamma)$, subject to the sample analog of (4.4)

$$\sum_{t=1}^{T}\mathbf{z}_t'\left(Y(k,t)-\sum_{j=1}^{K}Y(j,t-1)\,p(j,k,t)\right)=\mathbf{0}, \tag{4.6}$$

for each $j = 2, \ldots, K$ and the row-sum constraint

$$\sum_{k=1}^{K}p(j,k,t)=1, \tag{4.7}$$

over all *j* and *t.* Given a sample of indirect noisy observations and corresponding moment conditions, the parametric family of convex divergences defined in Section 4.1, may be used to choose an optimum member of the CR family.

4.3 Unlocking The Content of Economic Time Series

Over the centuries economists have been interested in the informational content of time series data. Many methods have been proposed for a proper analysis of the time series probability space, but their applicability depends on the particular characteristics of the data. In spite of the many productive efforts in this area, the hidden temporal patterns underlying the time-dated outcomes has in general, remained hidden. The function of dynamics in this process is to connect temporal information from the behavioral environment, to system outcomes later on in time. Given this background and a sample of economic time series data, information theoretic--Permutation Entropy methods are suggested to gain insights into the hidden temporal dynamics underlying the possible state(s) of an economic system that is seldom if ever in equilibrium.

The Permutation Entropy provides a basis for the analysis of nonlinear time series and permits a way of describing, in probability distribution form, the underlying dynamic state of the system. The objective is to extract qualitative information from a nonlinear times series in the form of temporal dynamics. The idea is based on permutation patterns-ordinal relations among values of a time series, that concerns temporal information derived from the dynamic properties of the economic system. The complexity of an economic system is thought of in terms of accessible patterns hidden in the process-system. If there is no pattern then the process is considered stochastic. To provide a probability distribution of the temporal dynamics that is linked to the sample space, Bandt and Pompe (2002), proposed a method that takes time causality into account by comparing time related observations-ordinal patterns, in a time series. They consider the order relation between time series instead of the individual values. Permutation patterns-partitions-vectors are developed, by comparing the order of neighboring observations.

The Bandt and Pompe (2002) methodology for evaluating the probability distribution *P* associated with the time series dynamical system under study, starts by considering partitions of the pertinent *D*-dimensional space that reveal relevant details of the ordinal structure of a given one-dimensional time series $S(t)=\{x_t; t=1,...,M\}$, with embedding dimension *D*>1 and time delay $\tau=1.$ Interest is on "ordinal patterns," of order *D*, generated by

$$(s)\mapsto\left(x_{s-(D-1)},x_{s-(D-2)},...,x_{s-1},x_s\right), \tag{4.8}$$

we assigns to each time s the D-dimensional vector of values at times $s, s-1, ..., s-(D-1)$. By "ordinal pattern" related to the time (s) we mean the permutation $\pi = (r_0, r_1, ..., r_{D-1})$ of $[0,1,...,D-1]$ defined by

$$x_{s-r_{D-1}} \leq x_{s-r_{D-2}} \leq ... \leq x_{s-r_1} \leq x_{s-r_0}. \tag{4.9}$$

Thus, for all the D! possible permutations π of order D, their associated relative frequencies can naturally be computed by the number of times this particular order sequence is found in the time series divided by the total number of sequences. The probability distribution $P = \{p(\pi)\}$ is defined by

$$p(\pi) = \frac{\#\{s \mid s \leq M - D + 1; (s), \text{ has type } \pi\}}{M - D + 1}. \tag{4.10}$$

In this expression, the symbol # stands for "number."

It is possible to create a probability distribution, whose elements are the frequencies associated with this i^{th} permutation pattern, i=1,2,…,D!. Using, for example, the gamma ≥ 0, with a uniform reference distribution, member of the CR family. The information content of such a distribution for the D! distinct assessable states may be defined as $PE = -\sum_{i=1}^{D!} p_i \ln p_i$, with normalized version $PE_{\text{norm}} = -\frac{1}{\ln D!}\sum_{i=1}^{D!} p_i \ln p_i$. For a review and applications of the Permutation Entropy method, see Zanin et al. (2012) and Kowalski et al. (2012).

4.4 Network Science Information Recovery

To go beyond traditional toy modeling and mathematical anomalies, a new network based paradigm is developing under the name of Network Science (for example see Willinger, Alderson and Doyle (2009) and Barabasi (2012), and the references contained therein). It is based on data sets that are indirect, incomplete and noisy. There are several things that make this approach attractive for information recovery in economics and the other social sciences.

In the economic-behavioral sciences, everything seems to depend on everything else and this fits right in to the interconnectedness-simultaneity of the nonlinear dynamic network paradigm. There is also a close link between evolving network structures and the equilibrium or disequilibrium of economic-behavioral systems. Finally, in terms of a methodology, network problems appear to be consistent with the Information Theoretic Approach to information recovery.

To indicate the applicability of information theoretic approach an example may be useful. In an economic-behavioral network the efficiency of information flow is predicated on discovering or designing protocols that efficiently route information. In many ways this is like a transportation network where the emphasis is on design and efficiency in routing the traffic flows (for example see, Castro et al. (2004) and the references therein). To carry this information flow analogy a bit farther, consider the problem of determining point-to-point traffic flows between sub networks when only aggregate origin-destinations traffic volumes are known. Given information about the network protocol in the form of a matrix A_{ij} that is composed of binary elements, traffic flows may be estimated from the noisy aggregate traffic data. Since the origin to

destination routes-unknowns are much larger than the origin-destination data, this results in an ill-posed linear inverse problem of the type first introduced in Section 3. If we write the inverse problem as

$$y_i = \sum_{j=1}^{N} A_{ij} p_j, \tag{4.11}$$

where y_i and A_{ij} are known and p_j is unknown and $\sum_{j=1}^{N} p_j = 1$, we may make use of the CR family of divergence measures (4.13) and write the problem as the following constrained optimization problem:

$$I(p,q,\gamma) \mid \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{N} \left(A_{ij} p_j - y_i \right) \sum_{j=1}^{n} p_j = 1, p_j \geq 0 \right). \tag{4.12}$$

This is just the solution to a standard problem when a function must be inferred from insufficient sample-data information. Thus network inference and monitoring problems have a strong resemblance to an inverse problem in which key aspects of a system are not directly observable. For an application of information theoretic methods to this type of network information flow problem, see Cho and Judge (2013).

5. SUMMING UP

Using traditional econometric methods to solve causal related stochastic inverse problems is, to use an old metaphor, akin to trying to make an omelet without breaking the eggs. Although a major desire for economic information is on causal influence recovery, the intersection of this objective, the use of indirect noisy effects data and the application of traditional direct econometric methods, leads in many

cases to an empty set. The data for the most part are observational, and the estimation-information recovery component requires methods designed for solving stochastic ill-posed inverse problems. Given the existence of invented equilibrium economic systems, indirect noisy economic data, and econometric tools ill suited to the task, moving economics in the direction of a probabilistic and predictive science is a goal yet to be achieved.

One possible solution is the use of information theoretic econometric methods that are designed to cope with stochastic inverse problems, and provide policy choices that are not drawings from a uniform-maximum entropy distribution. The implication is that unless we develop and apply econometric models and methods that are appropriate to the data and economic problems-questions at hand, our ability to understand and recover empirical system probabilities and to provide accurate predictions about dynamic economic processes, is going to continue to be limited and uninformed.

Much of the current instrumental variable econometric work concerned with policy effects and causal influence, is troubling for the following reasons:

i) Although economic processes and systems are, in general, simple in nature, they seldom if ever in equilibrium;

ii) The underlying hidden dynamics may be complicated, not well understood, and econometrically difficult to model;

iii) The usual family of possible subject matter econometric models may be incorrectly specified and contain certain specification inadequacies;

iv) The data come from an uncontrolled sampling process that is

observational in nature and may contain large systematic errors;

v) The sample observations consists of *mutual indirect noisy effects data*, *does not contain location or directional information and* makes it impossible to distinguish between *mutual influence and causal influence*;

vi) Even introducing a lag in the mutual indirect noisy observations fails to distinguish information that is actually exchanged from shared information, and does not support time causality and system dynamics;

vii) Combining conventional estimation and inference methods, with fragile subject matter econometric models, often leads to large parameter estimation biases and incorrect inferences;

viii) Confounders abound, since in economic processes and systems, everything seems to depend on everything else;

ix) These propositions taken either individually or in combination, dominates the statistical variation on which policy effects, causal influence and causality is based.

The above observations would seem to indicate that economic-econometric information recovery requires new modeling and econometric methods, if we are to understand the underlying dynamic economic systems that generate the indirect noisy effects data that we observe. Given sufficient parameterization it is always possible to match a model to any data set, without capturing the underlying hidden structure or the key functionality of the system at hand. Our diligent pursuit of measurement with current analytical economic models and estimation and inference methods, has in many cases

reduced traditional econometrics to data fitting and a not very useful vision of the economic world we seek to understand. Finally it is important to note that the internet has become a popular domain for the development of new theories related to the organization and behavior of large scale complex and dynamic models. In this context, the current emergence of noisy large scale data sets and network theory-modeling (e.g., Willinger et al. 2008), linked to information theoretic methods that appear to evoke some form of optimization(see Section 4.4), appears to offer promising new information recovery possibilities.

6. REFERENCES

Bandt, Christoph and Bernd Pompe. 2002. "Permutation Entropy: A Natural Complexity Measure for Time Series." *Physical Review Letters* 88(17)174102-1-174102-4.

Barabasi, Albert-Laszlo. 2012. "The Network Takeover." *Nature Physics* 8:14-16.

Caballero, Ricardo J. 2010. "Macroeconomics After The Crisis: Time To Deal With The Pretense-of-Knowledge Syndrome." NBER Working Paper.

Castro, Rui, Mark Coates, Gang Laing, Robert Nowak and Bin Yu. 2004. "Network Tomography: Recent Developments." *Statistical Science* 19:499-517.

Cho, Wendy and George G. Judge. ed. 2013. "An Information Theoretic Approach to Network Tomography." Working Paper, UC Berkeley.

Cressie, Noel and Timothy R.C. Read. 1984. "Multinomial Goodness-of-Fit Tests." *Journal of The Royal Statistical Society, Series B* 46:440–64.

Gorban, Alexander N. ed. 1984. "Equilibrium Encircling: Equations of Chemical Kinetics and Their Thermodynamic Analysis." Nauka: Novosibirsk.

Gorban, Alexander N. and Iliya V. Karlin. 2003. "Family of Additive Entropy Functions out of Thermodynamic Limit." *Physical Review E* 67:016104.

Gorban, Alexander N., Pavel A. Gorban and George G. Judge. 2010. "Entropy: The Markov Ordering Approach." *Entropy* 12:1145-1193.

Grendar, Marian and George G. Judge. 2009. "Asymptotic Equivalence of Empirical Likelihood and Bayesian MAP." *Annals of Statistics* 37:2445-2457.

Judge, George G. and Ron C. Mittelhammer. ed. 2011. *An Information Theoretic Approach to Econometrics*. Cambridge University Press.

Judge, George G. and Ron C. Mittelhammer. 2012. “Implications of the Cressie-Read Family of Additive Divergences for Information Recovery.” *Entropy* 14:2427-2438.

Kowalski, Andres M., Maria T. Martin, Angelo Plastino and George G. Judge. 2012. “On Extracting Probability Distribution Information From Time Series.” *Entropy* 14:1829-1841.

Miller, Douglas and George G. Judge. 2012. “Information Recovery in a Dynamic Statistical Model.” (Retrieved from SSRN: http://ssrn.com/abstract=2183398).

Willinger, Walter, David Alderson and John C. Doyle. 2008. “Mathematics and the Internet: A Source of Enormous Confusion And Great Potential.” *Journal of the American Mathematical Society* 56:586-599.

Zanin, Massimilliano, Luciano Zunino, Osvaldo A. Rosso and David Papo. 2012. “Permutation Entropy And Its Main Biomedical and Econophysics Applications: A Review.” *Entropy* 14:1553-1577.