Proceedings of the 6th Polish Symposium of Physics in Economy and Social Sciences (FENS2012), Gdańsk, Poland

A Stochastic Non-Homogeneous Constant Elasticity of Substitution Production Function as an Inverse Problem: A Non-Extensive Entropy Estimation Approach

S. Bwanakare*

University of Information Technology and Management in Rzeszów, Rzeszów, Poland

The document proposes a new entropy-based approach for estimating the parameters of nonlinear and complex models, i.e. those whose no transformation renders linear in parameters. Presently, for estimating such class of functions, various iterative technics like the Gauss–Newton algorithm are applied and completed by the least square methods approaches. Due to conceptual nature of such methods, definitely estimated functions are different from the original nonlinear one and the estimated values of parameters are in most of cases far from the true values. The proposed approach, being related to the statistical theory of information, is very different from those so far applied for that class of functions. To apply the approach, we select a stochastic non-homogeneous constant elasticity of substitution aggregated production function of the 27 EU countries which we estimate maximizing a non-extensive entropy model under consistency restrictions related to the constant elasticity of substitution model plus regular normality conditions. The procedure might be seen as an attempt to generalize the recent works (e.g. Golan et al. 1996) on entropy econometrics in the case of ergodic systems, related to the Gibbs–Shannon maximum entropy principle. Since this nonlinear constant elasticity of substitution estimated model contains four parameters in one equation and statistical observations are limited to twelve years, we have to deal with an inverse problem and the statistical distribution law of the data generating system is unknown. Because of the above reasons, our approach moves away from the normal Gaussian hypothesis to the more general Levy instable time (or space) processes characterized by long memory, complex correlation and by a convergence, in relative long range, to the attraction basin of the central theorem limit. In such a case, fractal properties may eventually exist and the *q non extensive parameter* could give us useful information. Thus, as already suggested, we will propose to solve for a stochastic inverse problem through the generalized minimum entropy divergence under the constant elasticity of substitution model and other normalization factor restrictions. At the end, an inferential confidence interval for parameters is proposed. The output parameters from entropy formalism represent the long-run state of the system in equilibrium, and so, their interpretation is slightly different from the "ceteris paribus" interpretation related to the classical econometrical modeling. The approach seems to produce very efficient parameters in comparison to those obtained from the classical iterative nonlinear method which will be presented, too.

DOI: 10.12693/APhysPolA.123.502

PACS: 89.65.Gh, 89.70.cf

1. Tsallis entropy and low frequency series econometric model

This document considers that the Tsallis entropy should remain, *even in the case of low frequency series*, a precious device for econometrical modeling since outputs provided by the Gibbs–Shannon entropy approach correspond to the Tsallis entropy limiting case of the Tsallis q-parameter equal to unity. Another maybe more pertinent argument in favor of applying the Tsallis non--extensive entropy approach could result from the fact that a number of complex phenomena involves the long range correlations which, in particular, can be seen when data are timely scaled-aggregated [1, 2]. This could probably be owing to interaction nature between the functional relationships describing the involved phenomena and the inheritance properties of a power law (PL), or can be depending on their non-linearity. Delimiting the threshold values for a PL transition towards the Gaussian structure (or to the exponential family law) as a function of data frequency level, is difficult since each phenomenon may display its own rate of convergence — if any, towards the central theorem limit attractor.

The next source of statistical concern may come from the systematic errors owing to the statistical data collecting and processing. Such a situation eventually could lead to tail queues distribution, too. Thus, a systematic applying of the Shannon–Gibbs entropy approach in the above cases — even on a basis of annual data — could be misleading and lead to instable solutions, in the above yet bad known situations. In reverse, since the non-extensive Tsallis entropy generalizes the exponential family law [3], the q-Tsallis entropy methodology fits well to the high or low frequency series. Furthermore, among the class of a few types of higher-order entropy estimators able to generalize the Gaussian law, the Tsallis non-extensive entropy presents the additional valuable quality of concavity — then stability, along the existence interval characterizing most of the real world phenomena. As far as the q-generalization of the Kullback–Leibler (K–L) rela-

*e-mail: sbwanakare@wsiz.rzeszow.pl

tive entropy index is concerned, *the latter conserves the same basic properties as the standard* K–L entropy and can be used for the same purpose [4].

Finally, as a consequence of the above reasoning, the q-Tsallis parameter presents an expected advantage of monitoring complexity of systems by measuring at what extent a given random phenomenon is far from the Gaussian benchmark. This can help, among other advantages, in drawing some attention on the quality of collected data or on the involved distribution.

2. Non-homogeneous constant elasticity of substitution production function

In 1961 Arrow, Chenery, Minhas and Solow (ACMS) [5, 6] have proposed a new mathematical function which simultaneously displays the property of homogeneity, constant elasticity of substitution (CES) between factors of production and of the possibility of differentiating elasticity of substitution for different industries, sectors or countries. In this article we treat the generic case of an aggregating function of production using the two classical factors: labor and capital. Let us recall its mathematical form

$$VA_t = \alpha \left[\delta K_t^{-\rho} + (1-\delta) L_t^{-\rho}\right]^{-\frac{\nu}{\rho}} \mathrm{e}^{\varepsilon_t} \tag{1a}$$

or one of its generalized forms as

$$Y = \left[\sum_{i=1}^{n} a_i^{(1-\rho)} X_i^{p}\right]^{-\frac{\nu}{\rho}} \tag{1b}$$

where

$$\rho = \frac{1-\tau^{\mathrm{e}}}{\tau^{\mathrm{e}}} \quad \text{with} \quad -1 \prec \rho \prec +\infty$$
$$\text{and} \quad 0 \prec \tau^{\mathrm{e}} \prec +\infty. \tag{2}$$

τ^{e} constant elasticity of substitution, ε_t stands for the random disturbance with unknown distribution. In Eq. (1a), α stands for the shift parameter; the parameter δ belongs to zero-one interval and represents the share (distribution) of the sold quantities of both distributed factors. The parameter ν reflects a degree of changing returns (VA) to scale. The higher the value of ρ, the higher the degree of substitution between factors. When this parameter converges to ∞ we are dealing with the case of perfectly complementary factors. The case of ρ converging to 0 suggests perfectly substitutable factors. The generalized form (1b) suggests a case of more than two inputs X_i.

Let us underscore an important connection between the CES function and a PL. In fact, by aggregating into one variable all X_i of Eq. (1), we get a generic case of a PL. Due to ubiquity of both functions, this sounds to bear potential implications on the economical ground. Parameter estimation of a CES model requires non-linear fitting techniques which are quite complicated and outputs less reliable. Since this nonlinear function has 4 unknown parameters — however with 3 degrees of freedom, we deal with an inverse problem. For reaching better solutions, we propose here to move away from the classical Gibbs–Kullback–Leibler cross-entropy formalism to the more general Tsallis non-extensive relative entropy.

3. q-Generalization of the K–L relative entropy and constraining problems

The Kullback–Leibler index of information divergence[†] [7] can be straightforwardly q-generalized as follows [4, 8]:

$$I_q(p, p^{(0)}) \equiv -\int \mathrm{d}x p(x) \ln_q \left(\frac{p^{(0)}(x)}{p(x)}\right)$$
$$= \int \mathrm{d}x p(x) \frac{\left[p(x)/p^{(0)}(x)\right]^{q-1} - 1}{q-1} \tag{3}$$

or

$$I_q(p, p^{(0)}) \equiv \sum p_i \frac{\left[p_i/p_i^{(0)}\right]^{q-1} - 1}{q-1} \tag{4}$$

in discrete case. $I_q(p, p^{(0)})$ stands for the traditional Kullback–Leibler index of information divergence between hypotheses p and $p^{(0)}$, provided that q is equal to unity. There exist two main versions of the Kullback–Leibler divergence (K–Ld) in the Tsallis statistics, namely the usual generalized K–Ld shown above and the generalized Bregman K–Ld [9]. Following Venkatesan and Plastino in [10], problems have been encountered in empirical implementation while trying to reconcile them. In their recent study, above authors have revealed interesting aspects concerning empirical research when the q-generalized cross-entropy is associated with constraining information. Following a recent literature [10, 11], the generalized Kullback–Leibler defined by Eq. (11) could be more consistent with expectations and constraints form proposed by Tsallis–Mendes–Plastino [9] known as the q-averages or the escort distribution[‡]:

$$\langle y_q \rangle = \sum_i \frac{p_i^q}{\sum_i p_i^q} y_i. \tag{5}$$

4. A generalized non-extensive entropy econometric model

This paragraph applies the results of Golan et al. [12] and some other authors, e.g. [6], who have proposed to reparameterize the moment constraints of the model, while argument in criterion function is already explained in probability form. We need to transform variables of constraining the generalized linear model into weight-probabilities over a support point space defining each of the original random parameter. For the clarity of notations, let us present details of the below reparameterized

[†]See for e.g. Kullback (1968) for rich definition of this index and its connection with Bayesian formalism.

[‡]However, for computational reasons, we have definitely opted in this document for applying the Curado–Tsallis (C–T) constraints [11] of the form: $\langle y_q \rangle = \sum_i p_i^q y_i$.

CES model. Let us then consider the next general linear model Y of the form

$$Y = X\beta + \varepsilon, \tag{6}$$

where β values are not necessarily constrained between 0 and 1. The variable ε is an unobservable disturbance term with a finite variance. As in classical econometrics, variable Y represents the system whose image must be recovered, and X stands for a vector of covariates related to the system by unknown parameters β with unobservable disturbance ε to be estimated through observable error components e. If we treat each β_k $(k = 1 \ldots K)$ as a discrete random variable with compact support [2] and $2 < M < \infty$ possible outcomes, then we can express β_k as

$$\beta_k = \sum_{m=1}^{M} p_{km} v_{km} \forall k \in K, \tag{7a}$$

where p_{km} is the probability of outcome v_{km} and the probabilities must be non-negative and sum up to one. Similarly, by treating each element e_i of e as a finite and discrete random variable with compact support and $2 < M < \infty$ possible outcomes centred around zero, we can express e_i as

$$e_i = \sum_{j=1\ldots J} r_{nj} z_{nj}, \tag{7b}$$

where r_n is the probability of outcome z_n on the support space j. We will use the commonly adopted index n, here and in the remaining mathematical formulations, to design the number of statistical observations. It is worth noticing that the term e can be fixed as percentage of explained variable as an a priori Bayesian hypothesis.

Posterior probabilities within the support space may display non-Gaussian distribution. Thus, we do not restrict the model to the conjugated distributions. The element v_{km} constitutes an a priori information provided by a researcher, while p_{km} is an unknown posterior (probability) whose value must be determined by solving a divergence entropy problem.

In matrix notation, let us rewrite $\beta = VP$, with $p_{km} \geq 0$ and $\sum_{k=1}^{K} \sum_{m>2\ldots M} p_{km} = 1$, where again K is the number of parameters to be estimated and M the number of data points on the support space. Also, let $e = rw$, with $r_{nj} \geq 0$ and $\sum_{n=1}^{N} \sum_{j>2\ldots J}^{J} r_{nj} = 1$ for N the number of observations and J the number of data points on the support space for the error term. Then, using the escort distribution in model moment definition, the Tsallis cross entropy econometric (MTEE) model can be stated as

$$\min H_q(a\|a^0, p\|p^0, r\|r^0)$$

$$\equiv \alpha \left[\sum a_{jm} \frac{[a_{jm}/ao_{jm}]^{q-1} - 1}{q-1} + \ldots \right.$$

$$\left. + \sum p_{km} \frac{[p_{km}/po_{km}]^{q-1} - 1}{q-1} \right]$$

$$+ \beta \sum r_{nj} \frac{[r_{nj}/ro_{nj}]^{q-1} - 1}{q-1}, \tag{8}$$

subject to

$$\ln VA = \ln \left(\sum_{j=1}^{J} g_j \frac{a_j^q}{\sum_{j=1}^{J} a_j^q} \right) - \frac{\sum_{h=1}^{H} v_h \frac{w_h^q}{\sum_{h=1}^{H} w_h^q}}{\left(\sum_{m=1}^{M} v_m \frac{p_m^q}{\sum_{m=1}^{M} p_m^q} \right)}$$

$$\times \ln \left(\sum_{i=1}^{I} \left(t_i \frac{b_i^q}{\sum_{i=1}^{I} b_i^q} \right) L^{\left(\sum_{m=1}^{M} z_m \frac{p_m^q}{\sum_{m=1}^{M} p_m^q} \right)} \right.$$

$$\left. + 1 - \sum_{i=1}^{I} (t_i b_i) K^{\left(\sum_{m=1}^{M} z_m \frac{p_m^q}{\sum_{m=1}^{M} p_m^q} \right)} \right)$$

$$+ \sum_{n=1}^{N} \sum_{j=1}^{J} z_{nj} \frac{r_j^q}{\sum_{j=1}^{J} r_j^q}, \tag{9}$$

$$\sum_{j>2\ldots M}^{J} a_j = 1, \qquad \sum_{m>2\ldots M}^{M} p_m = 1, \tag{10}$$

$$\sum_{i>2\ldots I}^{I} b_i = 1, \tag{11}$$

$$\sum_{h>2\ldots I}^{H} w_h = 1. \tag{12}$$

For the formal presentation reason, the criterion function (Eq. (8)) does not include probabilities w_h explaining the degree of economy changing to scale and b_i the parameter of distribution between factors. In order to improve the quality estimated parameter, the additional a priori information can be added to (8)–(12). In the case of a CES model, economic theory exists to helping to predict the sign value variation domain for each parameter. Then we get

$$0 \leq \alpha = Ga < \infty, \tag{13}$$

$$-1 \leq \rho = Zp \leq \infty, \tag{14}$$

$$0 \leq \delta = Tb \leq 1, \tag{15}$$

where α, ρ, δ in Eq. (1a) stand for the original, "before--reparameterization" parameters. G, Z, T stand for the above original parameter support space with corresponding weight-probabilities a, p, b defining output posteriors. The G, Z, T support spaces are included in a general support space V (Eq. (7a)) supporting all the parameters of the constraining equation system. The weights α, β introduced in the above dual objective function may exercise a significant impact on the model outputs through the Lagrange multipliers which transmits constraining information to the objective function.

5. Parameter confidence area

In this paragraph we will propose an inference information index $s(a_j)$ as an equivalent to a standard parameter error measure in the case of classical econometrics. Equivalent of determination coefficient R^2 will be proposed too under the entropy symbol $S(\mathrm{Pr})$. The departure point is that the maximum level of entropy-uncertainty is reached when the non-relevant information-moment constraints are enforced. This leads to a uniform distribution of probabilities over the k states of the system. As we add each piece of informative data in the form of a constraint, a departure from the uniform distribution will result, which means an uncertainty shrinkage. Thus, the value of below proposed $S(\mathrm{Pr})$ should reflect, for the whole model, a global departure from the maximum uncertainty.

Let us follow formulations in [12] and propose a normalized non-extensive entropy measure of $s(a_j)$ and $S(\mathrm{Pr})$. From the Tsallis entropy definition, $S_q > 0$, let us consider now all possible micro-states of the model. This number vary with the number of the support space data points i $(i = 1 \dots M)$ and the number of parameters of the model j $(j = 1 \dots J)$. Entropy S_q vanishes (for all q) in the case of $M = 1$; and for $M \succ 1$, $q > 0$, whenever one of the p_i $(i = 1 \dots M)$ occurrence equals unity, the remaining probabilities, of course, vanish. A global, absolute maximum of S_q (for all q) is obtained, in the case of a uniform distribution, i.e. when all $p_i = 1/M$. In such an instance we have for two both systems the maximum entropy equal to

$$S_q\left(a_j\right) = \left(M^{1-q} - 1\right)\left(1 - q\right)^{-1} \tag{16}$$

and

$$S_q(r) = \left(n^{1-q} - 1\right)\left(1 - q\right)^{-1}. \tag{17}$$

In Eq. (17) n varies with the number of the support space data points and the number of observations of the model. We propose below a normalized entropy index in which the numerator stands for the calculated entropy of the system and the denominator displays the highest maximum entropy as shown above (Eqs. (16) and (17)):

$$s\left(a_j\right) = p_{ij}\frac{\left(p_{ij}/p^0_{ij}\right)^{q-1} - 1}{q-1} \Big/ \left(M^{1-q} - 1\right)\left(1 - q\right)^{-1}$$
$$= p_{ij}\frac{\left(p_{ij}/p^0_{ij}\right)^{q-1} - 1}{1 - M^{1-q}}, \tag{18}$$

with j varying from 1 to J (the number of parameters of the system) and i belonging to M (the number of support space points), with $M > 2$. The total number micro-states is obtained by multiplying the number of model parameters J by the number of support space points M with $M > 2$. Then $s(a_j)$ reports precision on the estimated parameters. Equation (19) reflects the non-additivity Tsallis entropy property for two independent systems. The first term $S(p)$ is related to the parameter probability distribution and the second $S(r)$ to the error disturbance probability

$$S\left(\hat{\mathrm{Pr}}\right) = \left[S\left(\hat{p} + \hat{r}\right)\right]$$
$$= \left\{\left[S\left(\hat{p}\right) + S\left(\hat{r}\right)\right] + (1-q)S\left(\hat{p}\right)S\left(\hat{r}\right)\right\}, \tag{19}$$

where

$$S(P) = \sum\sum p_{ij}\frac{(p_{ij}/p^0_{ij})^{q-1} - 1}{q-1} \Big/ (M^{1-q} - 1),$$

and

$$S(r) = -\sum r_{nf}\frac{(r_{nf}/ro_{nf})^{q-1} - 1}{q-1} \Big/ n(1 - F^{1-q}).$$

$S(\hat{\mathrm{Pr}})$ is then the sum of the normalized entropies related to the parameters of the model $S(\hat{p})$, and to the disturbance term $S(\hat{r})$. Likewise, the latter value $S(\hat{r})$ is derived for all observations n, with F the number of the data points on the support space of the estimated probabilities r related to the error term. The values of these normalized entropy indexes $S(_{ij})$, $S(\hat{\mathrm{Pr}})$ vary between zero and one. Its values, near unity, indicate a poor informative variable — with higher entropy, while lower values are, in reverse, an indication of a better informative variable about the model. Both indexes fulfil the basic Fisher–Rao–Cramer information index properties, among which are continuity, symmetry, maximum and additivity.

6. Model outputs and discussion

This paragraph presents the model outputs in the case of the EU (27 countries) aggregated *value added* (VA_t) by the labour (L_t) and *capital* (K_t) *components*. The observed data cover twelve years period and are presented in Table I. We use a code General algebraic modelling system (GAMS) and the solver Minos5 to compute the model.

TABLE I

Aggregated value added and its components (mld euro) for 27 EU countries.

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
VA_{mld}	7.820	8.136	8.427	8.578	8.978	9.294	9.760	10.288	10.288	9.777	10.149	10.412
K_{ml}	3.287	3.420	3.550	3.641	3.860	4.020	4.265	4.528	4.476	4.105	4.336	4.444
L_{ml}	4.427	4.606	4.760	4.819	4.990	5.149	5.374	5.630	5.683	5.554	5.691	5.834

Source: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/

Figure 1 provides a comparison between the outputs from the cross-entropy (VA_{entrop}) and the nonlinear least squares (VA_{nlls}) regressions. For the results obtained in both cases, we note a standard error variation coefficients of CV of 0.06% and 11.5% for respectively the cross-entropy and the NLLS approaches. An index CV is obtained by dividing the standard error model disturbance by the average value of the dependent variable, i.e. the value added VA_t. The interval of interest is the one between unity (the Shannon entropy point) and 7/3 (the minimum square error point in this problem).

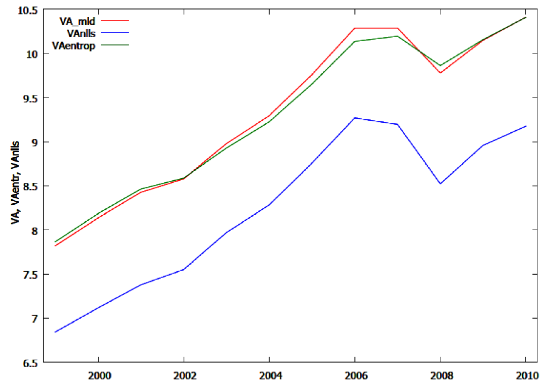

Fig. 1. Ex-post predictions of VA_t by q-entropy and NLLS approaches.

Figures 2 and 3 depict the relationships between the model error component CV and the q-Tsallis values, over a convex interval from 1 to 7/3 defining the CV minima for different values of q. The q-parameter has been incremented by a step of 0.5 starting from unity. Thus, this interval covers the Gaussian ($1 \prec q \prec 5/3$) and the partially stable laws (e.g. Levy's) attractors for ($5/3 \prec q \prec 3$). The purpose of below displayed figures is to depict the model disturbance structure dynamics for different q-values.

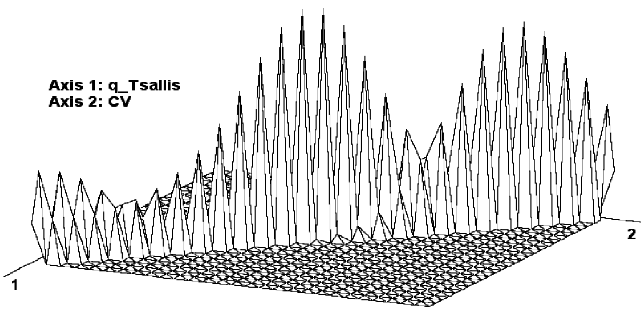

Fig. 2. Bivariate kernel density estimates between CV and q_Tsallis, for $[1.0 < q < 7/3]$.

Figure 2 displays the bivariate kernel joint density of $f(cv, q)$ with bandwidth $h = cn^{(-1/5)}$ where $c = 1$. Figure 2 is related to the interesting us interval of $[1 < q < 7/3]$. This higher bound is proposed because in this problem when q converges to 7/3 (see Fig. 3, point 25 on x-axis), CV reaches the global minimum over a convex space minimizing the considered criterion function. This corresponds to the best estimates of the model. We

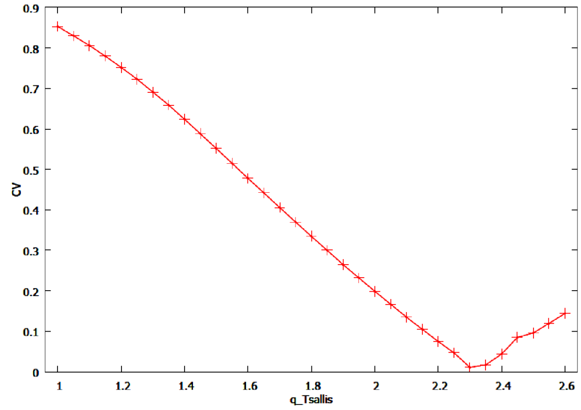

Fig. 3. Model disturbance (CV) curve as a function of q, for $[1 < q < 2.6]$.

would have expected solution for a q less than 5/3 for theoretical and empirical evidence.

However, comparable outputs may exist in literature. For example, Borges [13] found cumulative distribution of the scaled gross domestic product of the 167 countries around the world for the year 2000 corresponding to $q = 3.5$. This subject probably deserves further investigations. Coming back to Fig. 2, we observe 3 scaling structures of error components along with q-parameter steadily evolving towards the minimum point of the model error level.

Due to a low frequency twelve year period model sample, such a structure is most probably owing to the non-stationary series combined with the non-linear mathematical form of the CES function. As far as the model estimation is concerned, the model parameters have been initialized by the parameter output values from the nonlinear LS approach. The selected a priori parameter support space points for reparameterization vary between -5.5 and $+5.5$. The same prior space has been retained for the error disturbance but vary between -3 and $+3$, so as to conform it to the three sigma rule related to the Chebyshev inequality [14]. Both spaces are symmetric around zero, which prevents from bias of the estimated parameters. In the two component criterion function (Eq. (8)), we have retained a weight of 5% for the random disturbance. However, we have noticed the minimum error point of 7/3 to be less sensitive to weights.

6.1. *Parameter outputs of the Tsallis relative entropy model*

The parameters output of the Tsallis relative entropy model are given in Table II.

6.2. *Nonlinear LS estimation outputs*

In the case of the traditional nonlinear least square methods, Eq. (1) has first to be linearized using the Mac Lauren development, and next we apply the LS approaches [15] (Table III).

TABLE II

Dependent var: VA_t-aggregated EU value added.

Exogenous var: labour, capital	A	δ	p	v
estimates $_j$	1.866	0.163	0.001	1.0

Information index $I(S(\mathrm{Pr})) = 1 - S(\mathrm{Pr}) = 1 - 0.005 = 1.0$
variable q Tsallis parameter (for a weight $\alpha_i = 0.95\%$) $= 2.333$
$CV \approx 0.06\%$

The scaling parameter values A and the parameter v of changing returns (VA) to scale of both models are close to each other. However, since the error component is much higher in the case of the non-linear LS estimation, the q-cross entropy based estimates appear to be more reliable. Taking into account of the fact we deal with aggregated accounts of the 27 EU countries, the estimated parameters by the cross-entropy formalism remain conform to our expectations. The estimated parameter p with a value around zero suggests a convergence of the analyzed CES function to the well known by economists Cobb–Douglas function displaying, in the present case, constant returns to scale. A long-run optimal equilibrium share parameter δ between factors shows a lower proportion of labour of around 16.3% with respect to capital (83.7%). In 2010, this proportion was around 57% in favour of labour.

TABLE III

Dependent var: VA_t-aggregated EU value added.

Exogenous var: labour, capital	A	δ	p	v
estimates $_j$	1.995	0.282	3.046	0.993
standard error on model parameters (T-value)	48.89	6.61	1.49	6.61

coefficient of determination $R^2 \approx 0.88$
$CV \approx 11.5\%$

7. Concluding remarks

The present work has tried to develop a new Tsallis cross-entropy approach for econometric modelling, particularly in the case of inverse problem. The output values display a high precision and remain conform to our long run economic expectations. A global optimum of the model at very high value of q-parameter needs further verifications. Since the CES function stands for a particular form of a power law, the proposed approach should be suitable for future theoretical and empirical investigations on these classes of functions.

References

[1] S. Drożdż, J. Kwapień, *Phys. Rep.* **515**, 115 (2012).
[2] R. Rak, S. Drożdż, J. Kwapień, *Physica A: Statist. Mech. Appl.* **374**, 315 (2007).
[3] F. Nielsen, R. Nock, *J. Phys. A: Math. Theoret.* **45**, 032003 (2012).
[4] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, Berlin 2009.
[5] R. Klump, C. Papageorgiou, *J. Macroeconom.* **30**, 599 (2008).
[6] L. Secondi, *Entropy Approaches for the Production and Demand Function Parameter Estimation in a Regional CGE Model Framework*, www.ds.unifi.it/phd/dottorandi/tesi/TESI_LucaSecondi.pdf .
[7] S. Kullback, R.A. Leibler, *Ann. Math. Statist.* **22**, 79 (1951).
[8] M. Gell-Mann, C. Tsallis, *Nonextensive Entropy, Interdisciplinary Applications*, Oxford University Press, Oxford (USA) 2004.
[9] C. Tsallis, R.S. Mendes, A.R. Plastino, *Physica A: Statist. Mech. Appl.* **12**, 15 (1998).
[10] R.C. Venkatesan, A. Plastino, *Deformed Statistics Kullback-Leibler Divergence Minimization within a Scaled Bregman Framework*, arxiv/1102.1025v3, 2011.
[11] A. Sumiyoshi, G.B. Bagci, *Constraints and relative entropies in non-extensive statistical mechanics*, arxiv/cond-mat/0404253, 11 Apr 2004.
[12] A. Golan, G. Judge, D. Miller, *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Wiley, Chichester 1996.
[13] E.P. Borges, *Physica A* **334**, 255 (2004).
[14] F. Pukelsheim, *Am. Statist. Assoc.* **48**, 88 (1994).
[15] W.H. Green, *Econometrics Analysis*, 5th ed., Prentice Hall, New York 2003.