Proceedings of the 4th Polish Symposium on Econo- and Sociophysics, Rzeszów, Poland, May 7–9, 2009

Non-Extensive Entropy Econometric Model (NEE): the Case of Labour Demand in the Podkarpackie Province

S. Bwanakare*

University of Information Technology and Management (WSIZ), Rzeszów, Poland

The non-extensive entropy (NEE) principle has been successfully applied in the case of high frequency financial market analysis. I try to extend the approach to empirical social sciences and propose a competitive estimation approach with respect to classical econometrical methods. This article constitutes a limited extension of Jaynes–Shannon–Gibbs' (JSG) ergodic system formalism already applied to classical econometrics. The Podkarpackie private labour demand model is then developed and its outputs presented. A constrained weighted dual criterion function maximising entropy probabilities for parameter and disturbance components is derived and its inferential information indexes are proposed and computed. We note that the increase of relative weights on disturbance component leads to higher values of q, the entropic index of generalized Tsallis entropy. Smaller disturbance weights produce q values closer to unity. Outputs then converge to those displayed by the competitive JSG and least squares (LS) approaches. However, finding out an inferential rule delimiting the critical q values for Gaussian distribution interval remains of high interest. In terms of economics, the results of the proposed model show a realistic adjusting speed mechanism of actual lever of employment to its long run targeted equilibrium level owing to expected market profits.

PACS numbers: 89.65.Gh, 89.70.cf

1. Theoretical economic model

In this paragraph, I set up a labour demand model which focuses on the long run and short run relationships of labour demand determinants through an autoregressive self error correction process. In the short run, The Podkarpackie managers decide about the number of employees to be hired (or dismissed) in accordance with the long run expected optimal level of production. This usually happens in the case of expected shock in sales. However, because of institutional (for example trade union resistance) or economical reasons (for instance related to economic uncertainty), that optimal number is not hired (or dismissed) at once. In both cases the process of shock correction will be either short or long, depending on its origin and magnitude. Under neoclassical assumptions [6, p. 42], the desired level of labour demand L^* is a function of the output:

$$L^* = \alpha \exp^{-\beta t} Y. \tag{1}$$

Assuming that labour demand adjust to its targeted level by an error correct model:

$$\log\left(\frac{L_t}{L_{t-1}}\right) = \lambda \log\left(\frac{L_t^*}{L_{t-1}^*}\right) + \mu \log\left(\frac{L_t^*}{L_{t-1}}\right), \tag{2}$$

and combining (1) and (2) leads to:

$$\log\left(\frac{L_t}{L_{t-1}}\right) = \lambda \log\left(\frac{Y_t}{Y_{t-1}}\right) + \mu \log\left(\frac{Y_{t-1}}{L_{t-1}}\right) + \mu\beta t + \alpha o\,. \tag{3}$$

λ is the impact of output on labour demand and then an average long run elasticity of labour demand with respect to output Y_t. The coefficient $\mu$is the error correct parameter. Since a relation $0 \le \mu \le 1$ should prevail, this means that equilibrium error is only partly adjusted to each period. We note t as an index of discrete time, with $t = 1, 2, \ldots T$.

2. A generalized linear non extensive entropy econometric model

2.1. General model

This paragraph presents a generalized linear non extensive entropy econometric approach to estimate a model of Podkarpackie labour demand determinants. Following A. Golan [1], we first reparametrize the generalized linear model before fitting it to Eq. (3).

Let us consider the next general linear model:

$$Y = X \cdot \beta + e\,, \tag{4}$$

where β values are not necessarily constrained between 0 and 1 and ε is an unobservable disturbance term with finite variance, owing to economic data nature of exhibiting errors observation from empirical measurement or from random shocks. As in classical econometrics, variable Y represents the system which image must be recovered, and X accounts for explanatory variables generating the system with unobservable disturbance e. Unlike classical models, any constraining hypothesis is needed. In particular, the number of parameters to be estimated may be higher than observed data points [1]. Additionally, to increase precision of such parameters estimated from poor quality of data points, the entropy objective function allows for incorporation of all constraining functions which may act as a bayesian a priori information in the model.

* corresponding author; e-mail: `sbwanakare@wsiz.rzeszow.pl`

We treat each $\beta_k (k = 1, \ldots, K)$ as a discrete random variable with compact support and $2 < M < \infty$ possible outcomes. So, we can express β_k as

$$B_k = \sum_{m=1}^{\mathrm{M}} p_{km} v_{km} \quad \forall k \in K\,, \tag{5a}$$

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

$$e_i = \sum_{j=1.J} r_{nj} \cdot w_{nj}\,, \tag{5b}$$

where r_n is the probability of outcome w_n on the support space j. We will use the commonly adopted index n, here and in the remaining mathematical formulations, for the number of statistical observations. However, since we are dealing with a time series model, using index t in place of n would be acceptable. 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. The element v_{km} constitutes an prior information provided by the researcher, while p_{km} is an unknown probability whose value must be determined by solving a maximum 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 $\sum_{n=1}^{N} \sum_{j>2\ldots J} r_{nj} = 1$ for N the number of observations and J the number of data points on the support space for the error term. Then, the Maximum Tsallis Entropy Econometric (MTEE) estimator can be stated as:

$$\max\left[H_q(p;r)\right] = \Big\{\Big[1 - \sum_k \sum_m \alpha (p_{km})^q\Big] + \Big[1 - \sum_n \sum_j (1-\alpha)(r_{nj})^q\Big]\Big\}(q-1)^{-1}\,, \tag{6}$$

subject to

$$Y = X\beta + e = XpV + rw\,, \tag{7}$$

$$\sum_{k=1}^{K} \sum_{m>2\ldots M} p_{km} = 1\,, \tag{8}$$

$$\sum_{n=1}^{N} \sum_{j>2\ldots J} r_{nj} = 1\,, \tag{9}$$

where the real q, as previously stated, stands for the Tsallis parameter. Above $H_q(p,r)$ weighted dual criterion function is non-linear and measures the entropy in the model. The estimates of the parameters and residual are sensitive to the length and position of support intervals of β parameters. Because parameters of proposed model concern elasticity or error correct coefficients which values laying between 0 and 1, then following [1], the support space should be defined inside interval zero and one. Additionally, within the same interval support, the model estimates and their variances should be affected by the number of support values [1]. Increasing the number of point values inside support space leads to improving the a priori information about the system. The weights α and $(1-\alpha)$ are introduced into the above dual objective function. The first term "of precision" accounts for deviations of the estimated parameters from the prior (defined under support space). The second of "prediction ex post" accounts for an empirical error term as a difference between predicted and observed data values of the model. We will then notice (see Fig. 1) different values of q Tsallis parameter — an indicator of (non)-Gaussian data distribution.

2.2. Application to private labour demand for the Podkarpackie province

The above formulation (7) has been applied in order to reparametrize Eq. (3) of labour demand. According to the above, in order to improve estimated parameter quality, one can add, accordingly to economic theory predictions, an additional *a priori* restrictions to (6)–(9) as follows:

$$0 \leq \lambda = Vp^{\wedge} \leq 1\,, \tag{10}$$

$$0 \leq \mu = Vp^{\wedge} \leq 1\,, \tag{11}$$

$$-\infty \leq \beta = Vp^{\wedge} \leq 1\,, \tag{12}$$

here λ, μ, and β have the same meaning as in Eq. (3) and $p^{\wedge}$ is posterior probability.

2.3. Estimated parameter confidence area

In this paragraph we will propose the normalized Tsallis entropy coefficient $S(\hat{a}_k)$ as an equivalent to a standard error measure in the case of classical econometrics. Equivalent of determination coefficient R^2 will be proposed too under the entropy symbol $S(\hat{\mathrm{P}}\mathrm{r})$. The departure point is that the maximum level of entropy--uncertainty is reached when significant information--moment constraints are not enforced. This leads to an 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 a lowering in uncertainty. Thus, the value of below proposed $S(\hat{\mathrm{P}}\mathrm{r})$ reflects a global departure from the maximum uncertainty for the whole model. Without giving superfluous theoretical details, we follow formulations in [3] and propose a normalized non extensive entropy measure of $S(\hat{a}_k)$ and $S(\hat{\mathrm{P}}\mathrm{r})$.

>From Tsallis entropy definition $S_q > 0$. The latter vanishes (for all q) in the case $M = 1$; and for $M > 1$, $q > 0$, whenever one of the p_i equals unity, the remaining ones, of course, vanish. A global, absolute maximum of S_q (for all q) is obtained, in the case of equiprobability, i.e. when all $p_i = 1/M$. Note that we are interested, for our economic analysis, in q values laying inside the interval (1, 5/3). In such an instance we have for our two

systems:

$$S_q(p) = \left(M^{1-q} - 1\right)(1-q)^{-1} \quad (13)$$

and

$$S_q(r) = \left(n^{1-q} - 1\right)(1-q)^{-1} \quad (14)$$

In the limit when $q = 1$, relation (13) leads to the Boltzmann–Shannon expression (Tsallis, 2004).

I propose below an normalized entropy index in which numerator stands for the calculated entropy of the system, while denominator displays the maximum entropy of the system owing to equiprobability property:

$$S(\hat{a}_k) = -\Big[1 - \sum_k \sum_m (p_{km})^q\Big] \Big/ \Big[k(M^{1-q} - 1)\Big], \quad (15)$$

with k varying from 1 to K (number of parameters of the system) and m belonging to M (number of support space points), with $M > 2$. $S(\hat{a}_k)$ then reports the precision of estimated parameters. Equation (16) reflects the non--additivity property of Tsallis entropy for two probably independent systems, the one of parameter probability distribution and the second of error disturbance probability distribution (plausibly with quasi-Gaussian properties):

$$S(\hat{\mathrm{P}}\mathrm{r}) = \big[S(\hat{p} + \hat{r})\big] = \big\{\big[S(\hat{p}) + S(\hat{r})\big] + (1-q)S(\hat{p})S(\hat{r})\big\}, \quad (16)$$

where $S(\hat{p}) = -[1 - \sum_k \sum_m (p_{km})^q]/[k(M^{1-q} - 1)]$ and $S(\hat{r}) = -[(1 - \sum_n \sum_j r^q)]/[kn(J^{1-q} - 1)]$.

$S(\hat{\mathrm{P}}\mathrm{r})$ is then the sum of normalized entropy related to parameters of the model, $S(\hat{p})$, and to disturbance term $S(\hat{r})$. Likewise, the latter value $S(\hat{r})$ is derived for all observations n, with J the number of data points on the support space of estimated probabilities r related to error term.

The values of this normalized entropy indexes $S(\hat{a}_k)$, $S(\hat{\mathrm{P}}\mathrm{r})$ vary between zero and one. Its values, near to one, indicate a poor informative variable while lower values are an indication of better informative estimated parameter $\hat{a}_k$ about the model.

3. Data and model outputs

The statistical annual data [8] used to compute parameters of the dynamic model (Eq. (3)) come from GUS database (Polish Main Statistical Office) and cover the period of 1995–2006. Data from statistical survey are usually deemed to display systematic errors [9, p. 153] which should lead to tall tail of Gaussian distribution.

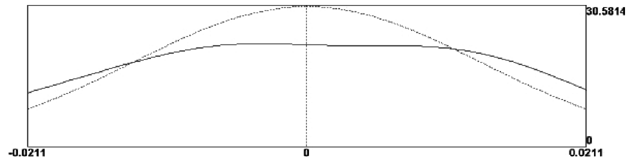

Fig. 1. Kernel density estimate of e_t as the OLS residuals of L_t/L_{t-1}.

Figure 1 tries to depict the distribution law of random components e_t of estimated model (outputs are below this page) in the presence of a very small data sample (eleven time series data points). Then, a kernel density estimate of e_t as the OLS residuals of L_t/L_{t-1} is displayed along with standard normal kernel density (dotted curve). We note a windows width $h = csT^{-0.2}$, with $c = 1$ and $s = 0.0130$, where s is the standard error of e_t. We observe that the depicted distribution displays a platykurtic shape, which may suggest a tall tail Gaussian distribution. The value level of parameter q could help to infer on some Gaussian distribution properties of the model.

Parameter estimation of entropy models has been computed using the solver PATHNLP incorporated in the software GAMS (General Algebraic Modelling System). We have noticed, through different simulations, that Shannon–Gibbs entropy model seems sensitive to initial conditions (support space of parameters in particular) that Tsallis entropy. This is a useful property particularly when an economic theory does not exist to prompt us which starting parameters to begin with.

Parameter estimation by robust standard errors' least squares (LS) approach has been carried out, using a freeware Gretl software. Thus, the HAC estimator is used for heteroscedasticity and autocorrelation correction.

3.1. Parameter outputs of Tsallis entropy model

Throughout many conducted experiments, we observe the coefficient $S(\hat{\mathrm{P}}\mathrm{r})$ to be very sensitive to weighting parameters α in the objective function.

TABLE

Dependent variable: $\log(L_t/L_{t-1})$.

Exogenous variables	$\log(y_t/y_{t-1})$	$\log(Y_{t-1}/L_{t-1})$	T	a_0
Estimates $\hat{a}_j$	0.168	0.495	−0.013	0.037
Precision error $S(\hat{a}_k)$ on estimated parameters	0.246	0.209	0.250	0.250

Information index $I[S(\hat{\mathrm{P}}\mathrm{r})] = 1 - S(\hat{\mathrm{P}}\mathrm{r}) = 0.782$

Variable q, Tsallis parameter (for a weight $\alpha_i = 9.1\%$) = 1.028

Average Time of Adjustment ATA [6] = $(1 - \lambda)/\mu \cong 1.68$ years

Tsallis q value being itself influenced by the above weights, its values closer or higher to 5/3 correspond to meaningless information index coefficients for which $S(\hat{\mathrm{Pr}})$ vanishes to zero. Inside the interval of parameter weights from 0.1 to 0.091, then q values sharply shift from 1.312 to 1.028. For next smaller values, optimal solution does not numerically exist (see Fig. 1). Interestingly enough, as attested by here displayed on this page solution models, at that lever of q equal 1.028, outputs from the three presented methods become practically similar.

3.2. Parameter outputs of Shannon–Gibbs entropy model

TABLE

Dependent variable: $\log(L_t/L_{t-1})$.

Exogenous variables	$\log(y_t/y_{t-1})$	$\log(Y_{t-1}/L_{t-1})$	T	a_0
Estimates $\hat{a}_j$	0.168	0.495	−0.013	0.037
Precision error $S(\hat{a}_k)$ on estimated parameters	0.231	0.197	0.235	0.235

Information index $I[S(\hat{\mathrm{Pr}})] = 1 - S(\hat{\mathrm{Pr}}) = 0.775$

3.3. Robust standard errors LS estimation

TABLE

Dependent variable: $\log(L_t/L_{t-1})$.

Exogenous variables	$\log(y_t/y_{t-1})$	$\log(Y_{t-1}/L_{t-1})$	T	a_0
Estimates $\hat{a}_j$	0.1679	0.4947	−0.0134	0.0370
P-values	(0.09214)	(0.00067)	(0.00006)	(0.03885)

Corrected $R^2 = 0.84 = 0.838332$, $DW = 1.30785$

Three parameters are different from zero at 5% and one on the variable $\log(y_t/y_{t-1})_1$ significant at 10%. The above good precision of the estimated parameters, from such a too small data sample of an autoregressive model, suggests the presence of co integrating — at the same order — variables L_t and Y_t. Such a particular situation leads to super consistency of estimated parameters [5].

Figure 2 presents the function $q = f(\alpha_i)$ with $(i = 1, \ldots, 417)$ which is concave.

The argument is the weights allocated to probabilities p_i in the criterion function. Values of $1 - \alpha$ higher than 0.909 do not provide an optimal solution. Then, $(1-\alpha) = 0.909$ corresponds to $q = 1.028$. Over the x-axis, we note the number of steps starting from one and repeated 415 times. The smallest weight α_i for which the criterion function is derivable starts with α_1 equal to 0.091, and its equivalent q value equal to 1.028 is reported on the y-axis. When $\alpha_i = 0.128$, the q value equals 5/3, as has already been explained. We stopped the experiment after 415 steps consisting in increasing α_i. Power law q is then equal to 2.9.

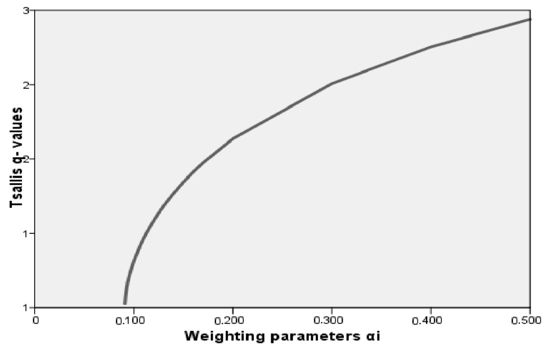

Fig. 2. Tsallis q-values as a function of weighting parameters' variation in the criterion function.

4. Comment on outputs and conclusions

In the present work we tried to develop a new Tsallis entropy approach for econometric modelling. For comparative reasons, two competitive econometric estimation methods have been present above. We note the accuracy in the similarity of outputs from the three models. We should remember that for the above entropy models, adding a new piece of a priori pertinent information could lead to better results.

On the economic side, we observe an average period of 1.68 years to adjust the actual lever of labour demand to its long-run targeted equilibrium owing to the expected market perspectives. Thus, entrepreneurs in the Podkarpackie province averagely require around one year plus 8 months to adjust the actual lever of labour demand to its long-run level which corresponds to the expected gross profits. The impact parameter is around 0.168. This is, on average, a 0.168% growth of labour demands when gross profits shifts up to 1%. As far as exogenous technical progress is concerned, we observe an expected negative sign on the value of the estimated parameter β, on the symptomatic variable t.

To conclude, in spite of positive results, much more research needs to be carried out to put in exergue the strong and weak sides of the proposed Tsallis entropy approach. In particular, finding out the critical q value which corresponds to Gaussian distribution interval would be of high interest.

References

[1] A. Golan, G. Judge, D. Miller, *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Wiley, 1996.

[2] M. Gell-Mann, C. Tsallis, *Nonextensive Entropy, Interdisciplinary Applications*, Oxford University Press, 2004.

[3] A. Golan, Jeffrey M. Perloff, *Comparison of Maximum Entropy and Higher-Order Entropy Estimators*, University of California Press, Berkeley, 2001.

[4] E.T. Jaynes, *Probability Theory: the Logic of Science*, Washington University Press, 1994.

[5] W.H. Green, *Econometrics Analysis*, V ed., Prentice Hall, N.Y, 2003.

[6] INSEE, *Structures et proprietés de cinq modèles macroéconomiques françaises*, Malakoff Cedex, France, 1996.

[7] A. Giffin, A. Caticha, *Updating Probabilities with Data and Moments*, Department of Physics, University at Albany-SUNY, 2007.

[8] *Polish Office of Statistics* (GUS), `http://www.stat.gov.pl/gus/` .

[9] A. Plastino, A.R. Plastino, Tsallis Entropy and Jaynes' Information Theory Formalism, *Brazilian Journal of Physics*, 1999.

[10] S. Bwanakare, WSIZ Rzeszow, Income Distribution and Fiscal Policy in Gabon: a Computable General Equilibrium Model with Maximum Entropy Principle, University of Gdańsk, Department of Management, 2007.