8. DISCOVERING THE SYSTEM STRUCTURE WITH APPLICATIONS IN ECONOMIC AND SOCIAL SCIENCES[1]

Ştefan ŞTEFĂNESCU[*]

Abstract

We develop Dobrescu's ideas (2011, 2013) to discover the main characteristics of an economic or social system. Many classifications of models, based on different dissimilarity measures, were taken into consideration. Also a practical example, which analyzes the relations among some indicators regarding the quality of life in Romania in 2010, was given.

Keywords: system structure, dissimilarity measures, structural levels, hierarchical clustering, quality of life indicators

JEL Classification: C18, C51, I32, P51

I. Introduction

The importance of the sectoral structure in the analysis of an economy for making a good prediction of economic growth is recognized (Dobrescu, 2011, 2013; Atkinson and Bourguignon 1982; Dijkman *et al.*; Erdem, 1996). Professor Emilian Dobrescu used in 2013 some similarity indices (structural coefficients) to measure the dependence relations among different economic sectors, finally intending to define some aggregate indicators. More precisely, Dobrescu measured in his paper Dobrescu, 2013 the similarity between the analyzed real structure and a control structure taken as referential to locate the changes in a dynamic economic system.

Our main aim is to apply Dobrescu's ideas to classify for the *QL* system some subjective indices which characterize the quality of life of a given population.

A first objection arises in this complex context. Since a similarity indicator between two objects is symmetric, we could not accurately indicate the tendency of evolution. In fact, the positive trend is a subjective convention imposed by an agreed point of view.

Moreover, by taking into account two antithetic referential points instead of a single one we improve the graphical accuracy regarding the system structure.

[1] Research Institute for the Quality of Life, Romanian Academy.
[*] E-mail: stefanst@fmi.unibuc.ro.

Of course, a referential variable is more or less subjective. For this reason, we sometimes prefer to classify the system components without using external information that is some inferential references. Such a kind of classification is based exclusively on the internal relations among the system variables. This approach will be exemplified in the paper, too.

II. Methodological Aspects

II.1. Dissimilarity Measures

To simplify the exposure, in the following we consider only the objects O having the form

$$O = \{ \underline{x} \mid \underline{x} = (x_1, x_2, x_3, ..., x_m), x_i \geq 0, 1 \leq i \leq m, x_1 + x_2 + x_3 + ... + x_m = 1 \}$$

The elements $\underline{x} \in O$ can be interpreted as cumulative distribution functions of some random variables.

In the set O we consider a proximity measure $\pi : O^2 \to O$, where $\pi(\underline{x}, \underline{y})$ appraises how close any two arbitrary objects $\underline{x}$ and $\underline{y}$ from O are. The proximity indicator π can be regarded as a dissimilarity measure $\pi_d(\underline{x}, \underline{y})$, that is how distinct the objects $\underline{x}$ and $\underline{y}$ are. In opposition, the similarity coefficient $\pi_s(\underline{x}, \underline{y})$ signifies how alike the elements $\underline{x}$ and $\underline{y}$ are.

A similarity measure satisfies the following axioms:

A1. $\pi_s(\underline{x}, \underline{y}) \leq \pi_s(\underline{x}, \underline{x})$ for any $\underline{x} \in O$, $\underline{y} \in O$ (an order between the objects)

A2. $\pi_s(\underline{x}, \underline{y}) = \pi_s(\underline{y}, \underline{x})$ for every $\underline{x} \in O$, $\underline{y} \in O$ (symmetry)

A3. $\pi_s(\underline{x}, \underline{y}) \geq 0$ for any $\underline{x} \in O$, $\underline{y} \in O$ (positive values)

A4. $\pi_s(\underline{x}, \underline{x}) = 1$ for any $\underline{x} \in O$ (normality property)

The last two assumptions are not always necessary in practice. More exactly, from a bounded indicator π which verifies the first two axioms we can build a new index satisfying all the assumptions A1-A4 (Everitt et. al., 2001; Fukunaga, 1990; Müller et. al., 2005).

In practice, it is important to choose the appropriate similarity measures from a set of such indices (Müller et. al., 2005). For this reason, we must have clearly in mind what "more similar" means for our concrete problem. In this context, the selected similarity indicator ought to also satisfy the following assumption:

A5. $\pi_s(\underline{x}, \underline{y}) > \pi_s(\underline{x}, \underline{z})$ if $\underline{x}$ is "more similar" to $\underline{y}$ than to $\underline{z}$

where $\underline{x} \in O$, $\underline{y} \in O$, $\underline{z} \in O$, $\underline{y} \neq \underline{z}$ (compatibility with reality)

We mention here how an appropriate similarity measure could be selected for particular types of biological data (Müller, Selinski, Ickstadt [11]) or to adjust a similarity indicator for objects having many attributes (Yalonetzky, 2010).

From a chosen similarity coefficient $\pi_s(\underline{x},\underline{y})$ we can obtain a lot of dissimilarity indices $\pi_d(\underline{x},\underline{y})$. So, if $h:[0,1]\to[0,1]$ is a strict decreasing function with $h(0)=1$ and $h(1)=0$ then we consider

$$\pi_d(\underline{x},\underline{y}) = h(\pi_s(\underline{x},\underline{y})) \ , \ \underline{x}\in O, \ \underline{y}\in O$$

When $h(s)=1-s$ we deduce a very useful dissimilarity coefficient $\pi_d(\underline{x},\underline{y})$, that is

$$\pi_d(\underline{x},\underline{y}) = 1-\pi_s(\underline{x},\underline{y})$$

Compared to a dissimilarity measure, a distance (metric) $\delta : O^2 \to O$ satisfies the restrictions

B1. $\delta(\underline{x},\underline{x})=0$ for all $\underline{x}\in O$ (lower bound for identical objects)

B2. For all $\underline{x}\in O$, $\underline{y}\in O$ with $\underline{x}\neq\underline{y}$ we have $\delta(\underline{x},\underline{y})>0$ (strictly positive)

B3. $\delta(\underline{x},\underline{y})=\delta(\underline{y},\underline{x})$ for any $\underline{x}\in O$, $\underline{y}\in O$ (symmetry)

B4. $\delta(\underline{x},\underline{y})\le\delta(\underline{x},\underline{z})+\delta(\underline{z},\underline{y})$ for every $\underline{x}\in O$, $\underline{y}\in O$, $\underline{z}\in O$ (triangle rule)

To underline the importance of the system structure when diverse economic models are investigated, Dobrescu studied in his paper [3] ten dissimilarity measures and he finally used five of them. Having in mind Dobrescu's comments from 2013, we prefer in our paper the following dissimilarity measures for the objects of the set O :

$$\delta_1(\underline{x},\underline{y}) = \sum_{i=1}^{m} |x_i - y_i| \qquad \text{(City block metric)}$$

$$\delta_2(\underline{x},\underline{y}) = \left(\sum_{i=1}^{m} (x_i - y_i)^2\right)^{1/2} \qquad \text{(Euclidean distance)}$$

$$\delta_3(\underline{x},\underline{y}) = 1-\left(\sum_{i=1}^{m} x_i y_i\right) \Big/ \left(\left(\sum_{i=1}^{m} x_i^2\right)\left(\sum_{i=1}^{m} y_i^2\right)\right)^{1/2} \qquad \text{(Angular coefficient)}$$

$$\delta_4(\underline{x},\underline{y}) = \sum_{i=1}^{m} |x_i - y_i| / |x_i + y_i| \qquad \text{(Canberra distance)}$$

with the convention $0/0=0$;

More details about the properties of different dissimilarity coefficients are given by Dijkman *et al.*, 2011 Everitt *et al.* 2001, Yalonetzky, 2010.

II.2. Classification of the Objects

The structure of the set O could be emphasized by putting together in the same cluster the similar objects and besides to establish the differences between these groups (Everitt *et al*., 2001; Jain, Murty and Flynn, 1999; Fukunaga, 1990).

Clustering methods are very useful in many decision-making economic analyses.

A hard classification procedure can be easily obtained by considering a single referential point $\underline{v} \in O$. For example, we deduce the clusters “low”, “middle” or “high” which group together all the elements $\underline{x} \in O$ depending on how far away the object $\underline{x}$ is from the reference point $\underline{v}$, that is on the value of the distance $\delta(\underline{x},\underline{v})$. The use of a single dissimilarity coefficient $\delta(\underline{x},\underline{v})$ to specify the real position of $\underline{x}$ from $\underline{v}$ could generate serious interpretation errors. For example, the positive value $\delta(\underline{x},\underline{v})$ does not inform us on a one-dimensional scale if the point $\underline{x}$ is on the left or on the right side of the control element $\underline{v}$. In fact, the positive sense on a given direction is a subjective convention.

By increasing the number of the referential points we obtain a more accurate classification of the objects $\underline{x} \in O$. For example, if we use two control variables $\underline{v}$ and $\underline{w}$ then every object $\underline{x} \in O$ has a bidimensional representation given by the rectangular coordinates $(\delta(\underline{x},\underline{v}),\delta(\underline{x},\underline{v}))$. Taking into consideration the distances among all these points $\underline{x} \in O$ we can create diverse groups.

But the usage of referential variables does not take into consideration the inner dissimilarities $\delta(\underline{x},\underline{y})$ which exist among all objects $\underline{x} \in O$, $\underline{y} \in O$. Having in mind this major disadvantage, we suggest a new approach based on a classification procedure of the elements from the system O.

More precisely, two arbitrary clusters are merged together to form a new larger cluster. The selection of the suitable pair of clusters to be merged is based on the minimum distance criteria.

The single-link algorithm uses the dissimilarity measure $\delta^{-}(A,B)$ between any two groups A and B, where

$$\delta^{-}(A,B) = minimum \; \{ \delta(\underline{x},\underline{y}) \mid \underline{x} \in A, \underline{y} \in B\}$$

$\delta(\underline{x},\underline{y})$ being a chosen dissimilarity function among the individuals $\underline{x}$ and $\underline{y}$.

By contrast, the complete-link procedure is based on an opposite dissimilarity index $\delta^{+}(A,B)$ for the clusters A and B, that is

$$\delta^{+}(A,B) = maximum \; \{ \delta(\underline{x},\underline{y}) \mid \underline{x} \in A, \underline{y} \in B\}$$

In the literature there is proof that the single-link algorithm produces elongated clusters in opposition to the complete-link method, which in general gives compact groups.

Having in mind all these comments, the hierarchical agglomerative clustering ***HAC*** procedure has the form (see also Everitt *et al.*, 2001; Jain *et al.*, 1999) :

Algorithm HAC.

Step 1. Initially, each object $\underline{x}$ forms its own cluster with a single element

Specify a dissimilarity measure $d^*(A,B)$ for any two clusters A and B.

Step 2. Compute the dissimilarities $d^*(A,B)$ for any pair (A,B) of clusters.

Step 3. Find the most similar pair (C,D) of clusters, that is

$$d^*(C,D) = \underset{A,B}{minimum} \{ d^*(A,B) \}$$

Step 4. Merge the most alike clusters A and B into a single new cluster.

Step 5. Repeat *Steps* 2-4 to obtain finally only a single cluster.

Remark. The "distance" $d^*(A,B)$ could be one of the previous dissimilarity coefficients $d^-(A,B)$ or $d^+(A,B)$. In the literature, more types of such dissimilarity indices between groups are analyzed (Everitt *et al.*, 2001).

A dendrogram represents graphically the nested grouping of clusters and at the same time establishes the levels at which new groups appeared (Everitt *et al.*, 2001; Jain *et al.*, 1999).

II.3. Dimension Reduction

In fact, every object $\underline{x} \in O$ is characterized by m attributes.

Since $\underline{x} = (x_1, x_2, x_3, \ldots, x_m)$ with $x_1 + x_2 + x_3 + \ldots + x_m = 1$ and $x_i \geq 0$, $1 \leq i \leq m$, then an exact representation of the point $\underline{x}$ is obtained in a real space with $m-1$ dimensions. When $m \geq 4$, the graphical image of the elements of the set O is very difficult to be interpreted.

In this situation, it is preferable to find for every element $\underline{x} \in O$ a proxy point $\underline{x}^{\#}$, $\underline{x}^{\#} = (x_1^{\#}, x_2^{\#})$ such that

$$\delta(\underline{x}, \underline{y}) \approx \delta(\underline{x}^{\#}, \underline{y}^{\#}) \text{ for all } \underline{x} \in O,\ \underline{y} \in O.$$

The bidimensional coordinates $x_1^{\#}, x_2^{\#}$ of the points $\underline{x}^{\#}$ are deduced by minimizing the expression

$$h_1(x_1^{\#},x_2^{\#},y_1^{\#},y_2^{\#},...)=\sum_{\underline{z},\underline{t}}(\delta(\underline{z},\underline{t})-\delta((z_1^{\#},z_2^{\#}),(t_1^{\#},t_2^{\#})))^2$$

The advantage to work with the points $\underline{x}^{\#}$ is that they are representable in a bidimensional space.

An efficient procedure to obtain the coordinates $x_1^{\#},x_2^{\#}$ of the points $\underline{x}^{\#}$ is based on the singular value decomposition method (Golub and Van Loan, 1996).

The coefficients γ_1 and γ_2 give the mean of the absolute and relative errors when the objects $\underline{x}\in O$ are approximated by the elements $\underline{x}^{\#}$. More precisely,

$$\gamma_1=\frac{1}{k(k-1)}\sum_{\underline{x}\neq\underline{y}}|\,\delta(\underline{x},\underline{y})-\delta(\underline{x}^{\#},\underline{y}^{\#})\,|$$

$$\gamma_2=\frac{1}{k(k-1)}\sum_{\underline{x}\neq\underline{y}}|\,\delta(\underline{x},\underline{y})-\delta(\underline{x}^{\#},\underline{y}^{\#})\,|\,/\,\delta(\underline{x},\underline{y})$$

where: k is the number of components from the system O.

II.4. Ordering the Objects

The dissimilarity measure does not necessarily imply a transitivity type property. For example, inequalities $\delta(\underline{x},\underline{y})\geq a$ and $\delta(\underline{y},\underline{z})\geq a$ do not necessarily imply the relation $\delta(\underline{x},\underline{z})\geq a$.

A binary relation "$\lhd$" between the objects $\underline{x}\in O$ is a relation of order if the following conditions are satisfied:

C1. For any $\underline{x}\in O$ we have $\underline{x}\lhd\underline{x}$ (reflexivity)

C2. If $\underline{x}\lhd\underline{y}$ and $\underline{y}\lhd\underline{x}$ then $\underline{y}=\underline{x}$ (anti-symmetry)

C3. If $\underline{x}\lhd\underline{y}$ and $\underline{y}\lhd\underline{z}$ then $\underline{x}\lhd\underline{z}$ (transitivity)

The relation of order "$\lhd$" is *total* if in addition we have the property

C4. For any $\underline{x}\in O$ and $\underline{y}\in O$ we have $\underline{x}\lhd\underline{y}$ or $\underline{y}\lhd\underline{x}$ (any two elements of the system O are comparable).

The binary relation "$\lhd$" defines a *partial order* if the condition **C4** is not true, as there are at least two noncomparable objects $\underline{x}\in O$, $\underline{y}\in O$ (we have neither $\underline{x}\lhd\underline{y}$ nor $\underline{y}\lhd\underline{x}$).

Many partial ordering measures could be defined on the set O of the simple discrete random variables (see, for instance, Giovagnoli and Wynn, 2008). In the following, we work with the partial stochastic ordering relation "$\prec$" for the elements of the set O.

More precisely, for any $\underline{x} \in O$, $\underline{y} \in O$ we have $\underline{x} \prec \underline{y}$ if and only if

$$\sum_{i=1}^{k} x_i \geq \sum_{i=1}^{k} y_i \quad \text{, for any } 1 \leq k \leq m .$$

We also use the notation $\underline{y} \succ \underline{x}$ when $\underline{x} \prec \underline{y}$.

In the case $\underline{x} \prec \underline{y}$ we say that $\underline{y}$ dominates $\underline{x}$ stochastically. Thus, for any simple discrete random variables X,Y which take only the values $1,2,3,\ldots,m$ with the mass probabilities $\underline{x}$, and $\underline{y}$, respectively, the relation $\underline{x} \prec \underline{y}$ necessarily implies the inequality $Mean(X) \leq Mean(Y)$.

The presence of a partial order relation "$\prec$" in the set O imposes a special level structure inside O. More exactly, any two objects $\underline{x},\underline{y}$ which belong to the same level L_j are noncomparable. In addition, between the elements of two consecutive levels L_j and L_{j+1} a "direct communication" is established, that is for any $\underline{y} \in L_{j+1}$ there is a $\underline{x} \in L_j$ with $\underline{x} \prec \underline{y}$ and we have not $\underline{z} \in O$, $\underline{z} \neq \underline{x}$ and $\underline{z} \neq \underline{y}$, such that $\underline{x} \prec \underline{z}$ and $\underline{z} \prec \underline{y}$.

III. A Practical Example

III.1. System Components

Table 1

The Meaning of the Indicators *A-N* (Questionnaire QL2010)

Name	Significance
A	Individual health
B	Family relations
C	Individual household
D	The quality of the environment
E	Work conditions
F	Relations with your neighbours
G	Family income
H	Access to drink water in your community
I	Health services received in your community
J	The police activity in your community

Name	Significance
K	The quality of the education in your community
L	The information received through mass media (press, radio, television)
M	The quality of public transport in your community
N	The current possibility to spend your free time (recreation facilities)

In the following, we intend to establish the interactions between the *A-N* components of the system *QL* which characterizes the quality of life.

The significance of the *A-N* variables is given in Table 1.

III.2. Data

Our statistical study uses data from a national representative sample ***E*** which was designed in 2010 by the Research Institute for the Quality of Life of the Romanian Academy (details regarding the database in [13]).

For every *A-N* question, a person chooses one of the following variants:

R1. “Very bad” – code 1; **R2.** “Bad” - code 2; **R3. “**Satisfactory” - code 3;

R4. “Good” - code 4; **R5.** “Very good” - code 5.

Table 2 shows the probabilities for the **R1-R5** answers taking into consideration 1045 individuals that have the Romanian nationality in the ***E*** sample ([13]).

In conclusion, the system *QL* which defines the quality of life for the Romanian people contains the *A-N* variables. Each of these variables is characterized by the distribution of the answers **R1-R5** (Table 2).

Table 2

The Distributions of the A-N Variables

Variable	R1	R2	R3	R4	R5	Total
A	0.082	0.170	0.281	0.386	0.081	1.000
B	0.008	0.011	0.100	0.613	0.268	1.000
C	0.011	0.053	0.216	0.594	0.126	1.000
D	0.011	0.095	0.285	0.514	0.095	1.000
E	0.034	0.117	0.313	0.447	0.089	1.000
F	0.007	0.016	0.126	0.668	0.183	1.000
G	0.152	0.220	0.396	0.217	0.015	1.000
H	0.054	0.139	0.195	0.527	0.085	1.000
I	0.042	0.117	0.289	0.503	0.049	1.000
J	0.025	0.086	0.343	0.509	0.037	1.000
K	0.014	0.094	0.307	0.535	0.050	1.000
L	0.014	0.082	0.257	0.558	0.089	1.000
M	0.027	0.073	0.267	0.564	0.069	1.000
N	0.063	0.204	0.366	0.317	0.050	1.000

III.3. Variables of Reference

We selected the dissimilarity measures $\delta_1 - \delta_4$ defined in section II.1.

We consider two referential variables, denoted by *V* and *W*. In fact, respecting the previous notations $V = (0.5, 0.5, 0.0, 0.0, 0.0)$, $W = (0.0, 0.0, 0.0, 0.5, 0.5)$ express a very bad, and a very good situation, respectively.

We attach to every *A*-*N* variable its rank resulted by reordering in an ascendant way the distances $\delta(X, Q)$, where $X \in \{A, B, ..., M, N\}$, *Q* being a referential point.

Table 3 synthesizes the scores received for the *A*-*N* variables. For a fixed referential point these scores do not depend very much on the dissimilarity measures used. Moreover, considering instead of the referential *V* its opposed point *W,* roughly a reverse order of the variable *A*-*N* is obtained (see the results in Table 3).

We also observe a special hierarchy location for variables *B*, *F*, and *G*, *N,* respectively.

Table 3

Ascending Order of the *X* Variables Based on the Distance $\delta(X, Q)$ to a Referential Point Q

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\delta_1(X,V)$	G	N	A	H	I	E	J	K	D	M	L	C	F	B
$\delta_2(X,V)$	G	N	A	H	E	I	D	J	K	L	M	C	B	F
$\delta_3(X,V)$	G	N	A	H	I	E	J	D	K	M	L	C	F	B
$\delta_4(X,V)$	G	N	A	H	I	E	J	K	D	M	L	C	F	B
$\delta_1(X,W)$	B	F	C	D	L	H	M	K	I	J	E	A	N	G
$\delta_2(X,W)$	B	F	C	H	L	D	M	E	I	A	K	J	N	G
$\delta_3(X,W)$	B	F	C	H	L	D	M	E	K	I	A	J	N	G
$\delta_4(X,W)$	B	F	C	D	H	L	E	M	I	A	K	J	N	G

III.4. A Bidimensional Image of the QL System

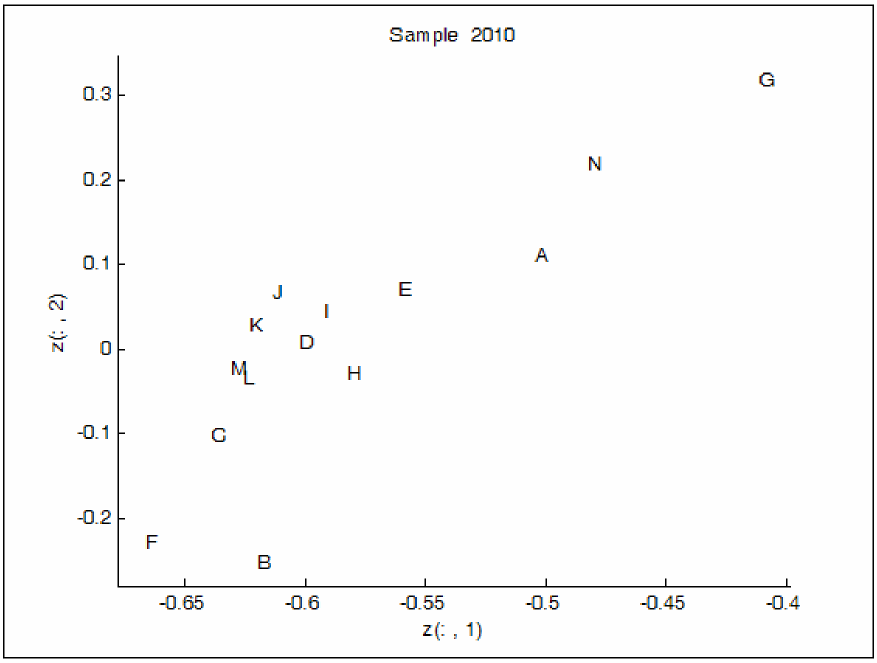
Applying a dimension reduction method (section II.3) the m-dimensional variable $\underline{x} \in O$ is approximated by a bidimensional vector $\underline{x}^{\#} = (x_1^{\#}, x_2^{\#})$. We denote this transform by Z, that is $Z(\underline{x}, 1) = x_1^{\#}$, $Z(\underline{x}, 2) = x_2^{\#}$.

Figure 1 suggests the position for all *A*-*N* components of the *QL* system respecting the real Euclidean distances between the elements. We notice here the singular point *G* (family income) and in the opposite side the variables *B* and *F* (the relations inside a family or with the neighbours).

The approximation procedure is very good, since the means of the absolute or relative errors γ_1, γ_2 are very small - more exactly, $\gamma_1 = 0.0195$ and $\gamma_2 = 0.1467$.

Figure 1

The "Distances" between the *A-N* Components of the *QL* System



III.5. Clustering

Table 5

Clustering with Three Classes

Method	Distance	Cluster 1	Cluster 2	Cluster 3
single	δ_1	G	B,F	A,C,D,E,H,I,J,K,L,M,N
single	δ_2	G	B,F	A,C,D,E,H,I,J,K,L,M,N
single	δ_3	G	N	A,B,C,D,E,F,H,I,J,K,L,M
single	δ_4	G	B,F	A,C,D,E,H,I,J,K,L,M,N
complete	δ_1	A,G,N	B,F	C,D,E,H,I,J,K,L,M
complete	δ_2	A,G,N	B,F	C,D,E,H,I,J,K,L,M
complete	δ_3	A,G,N	B,F	C,D,E,H,I,J,K,L,M
complete	δ_4	G	B,F	A,C,D,E,H,I,J,K,L,M,N

The distances among the *A-N* objects defined in Table 2 were computed independently using all the dissimilarity measures $\delta_1 - \delta_4$. Then, we applied the hierarchical agglomerative clustering procedure ***HAC*** with the single or complete link algorithm (Everitt *et al*., 2001; Jain *et al*., 1999). Considering for the classification procedure only three final clusters, Table 5 shows the sharing of the *A-N* components in every group. As a rule, a separate group contains the variable *G* (family income) and another distinct group has the *B, F* elements (the relations between persons). This characteristic of the *QL* system was already detected by both previous models (see Table 3 or Figure 1).

III.6. System Levels

Taking into consideration the relation of order "$\prec$" defined in section II.4 for the *QL* system characterized in Table 2 we finally obtained six interaction levels. Section II.4 presents the properties of the levels from an arbitrary system O.

Table 6 indicates the elements of each level. We notice again the *G* variable (family income) which belongs to the lower level and the *B, F* variables (family relations and the relationships with the neighbours) at the top level. All the previous models mentioned this opposite aspect. In addition, the classification of *QL* components by ascending levels clarifies the set of vulnerable variables.

Table 6

The Levels of the QL System

Level	1	2	3	4	5	6
Variables	G	A, I, J, N	E, H, K, M	D, L	C	B, F

IV. Concluding Remarks

More techniques which could be applied successfully in economic and social sciences to discover the inner structure of a given system were suggested.

The use of similar or different statistical models imposed the same final conclusions regarding the structure of the *QL* system which analyzes the quality of life in Romania in 2010. The selection mode of a dissimilarity measure does not affect decisively the results, the ratios of different elements of the *QL* system remaining almost unchanged.

To observe the constructive development of a system it is necessary to take into consideration both the "positive" and the "negative" points of reference (denoted in Table 3 by *W*, and *V,* respectively). In this way, we can establish correctly the real position of the entire system conditional on the exogenous variables *V* and *W*.

For the individuals having Romanian nationality, we noticed clearly in the *QL* "quality of life" system two opposed and very distinct components, namely $\{G\}$, and $\{B,F\}$ indices, respectively. We observe here a "materialism" aspect given by the variable G (family income) and also the sentimental size of the character suggested by the composite indicator $B+F$ (relations with the family and with the neighbours).

All the four investigation methods led for Romanian individuals to a very low subjective score regarding the "family income" and, at the same time, a very high appreciation concerning the relations with other people. We neglected deliberately an extensive discussion regarding all the *A-N* indicators that characterize the quality of life aspects.

Both papers of Professor Emilian Dobrescu (2011, 2013) revealed the importance of the structural coefficients (dissimilarity indices) to establish the ratios of different economic sectors. The present paper extended Dobrescu, 2013 ideas to discover the inner structure of an economic system.

References

Atkinson, A. and Bourguignon, F., 1982. The comparison of multi-dimensioned distributions of economic status. *Review of Economic Studies*, XLIX, pp.183-201.

Dijkman, R. et. al., 2011. Similarity of business process model : Metrics and evaluation. *Information Systems*, 36(2), pp.498-516.

Dobrescu, E., 2011. Sectoral structure and economic growth. *Romanian Journal of Economic Forecasting*, 14(3), pp.5-36.

Dobrescu, E., 2013. Restatement of the I-O coefficient stability problem. *Journal of Economic Structures*, 2(2), pp.1-84. Available at: <http://www.journalofeconomicstructures.com/content/2/1/2>.

Erdem, T., 1996. A dynamic analysis of market structure based on panel data. *Marketing Science*, 15, pp.359-370.

Everitt, B.S., Landau, S. and Leese, M., 2001. *Cluster analysis.* London: Arnold Publishers, Fourth edition.

Giovagnoli, A. and Wynn, H.P., 2008. Stochastic orderings for discrete random variables. *Statistics and Probability Letters*, 78, pp.2827-2835.

Golub, G.H. and Van Loan, C.F., 1996. *Matrix computations.* The John Hopkins University Press, Baltimore, Third edition.

Jain, A.K. Murty, M.N. and Flynn, P.J., 1999. Data clustering: A review. *ACM Computing Surveys*, 31(3), pp.264-323.

Fukunaga, K., 1990. *Statistical pattern recognition*. New York: Academic Press.

Müller, T. Selinski, S. and Ickstadt, K., 2005. Cluster analysis: A comparison of different similarity measures for SNP data. *Econstor*, Technical Report, Universität Dortmund, no. 2005-14, Available at: <http://hdl.handle.net/10419/22605>

Yalonetzky, G., 2010. A dissimilarity index of multidimensional inequality of opportunity. *Oxford Poverty & Human Development Initiative (OPHI)*, University of Oxford, Working paper no. 39.

***** 2010. *Data base DQL2010* (*Diagnosis for the quality of life–Romania 2010*). Research Institute for the Quality of Life, Romanian Academy.