ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 196, 2006

Tomasz Jurkiewicz, Ewa Wycinka***

EVALUATION OF THE DEGREE OF INTEGRATION AMONG EUROPEAN INSURANCE MARKETS

Abstract. The insurance internal market has existed since 1993 (enforced by the third life directive). Its' main features are a common framework to allow insurers to operate throughout the EU and to establish and provide services freely. On the other hand, the legal framework was designed to protect customers, particularly individuals, where the safe delivery of promised benefits can be vital.

One can observe that these frameworks do not guarantee that insurance markets develop in the same way. Insurers are more likely to set up their businesses in some countries than in others. There are also differences in: the number of policies, the amount of benefits and other indicators among European countries. The question is, whether we can talk about the internal market or rather a group of different national markets. Trying to answer this question, we have to take into account a set of variables that shows all the major aspects of integration.

The main purpose of this study is evaluating the level of European markets integration by using multivariate statistical methods. We shall also compare results obtained owing to application of different methods and will try to explain similarities and differences between the obtained results.

Key words: insurance markets, integration, multivariate analysis.

1. ORIGINS OF THE INTERNAL MARKET

The establishment of internal market was planed in the Treaty of Rome in 1957. The changes in law have been done in evaluative way to create insurance market more and more liberal (Śwital 1997). As the result of that process, insurance internal market has been existing since 1993. One of its' main features is a freedom of establishment and freedom of services through the whole Europe. It has been expected that these freedoms would allow to develop national markets in the same way. Thus, as the result

* Ph.D., Department of Statistics, University of Gdańsk.
** M.Sc., Department of Statistics, University of Gdańsk.

there would be one European internal market. However, one can easily find many factors that are incentives for companies to establish in some countries rather than in others. The most important factors are differences in national law and flexibility of regulatory authorities.[1]

Disproportions between markets can be noticed in: number of policies, amount of benefits and other indicators. There is a question whether we can still say about a single market or rather a set of different national markets. Moreover, we can ponder the similarities between particular markets and try to find clusters more homogenous than the European insurance market as the whole. To solve that problems we have used some statistical methods. We have decided to use a set of variables that shows all the aspects of integration rather than to focus on a single selected indicator.

2. DIAGNOSTIC VARIABLES

Insurance markets are described by wide range of variables presented as relative or absolute measures. Using only relative measures (structure ratio and intensity ratio) allows making comparison between small and large markets and the results are not directed in the largest markets. To eliminate the impact of different orders of variables, variables should be standardized.

We have decided to choose following variables:

• first group – indicated as *A* – are rates showing the importance of the insurance markets in economy as:[2]

*A*1 – premium/GDP,

*A*2 – investments/GDP,

*A*3 – investments in shares/market capitalization,

*A*4 – insurance employment/service,

*A*5 – premium per inhabitant,

*A*6 – insurance employment per inhabitant.

• the *B* group is composed of variables describing a structure of insurance markets, such as:

*B*1 – provisions/premium ratio,

*B*2 – number of companies per thousand inhabitants,

*B*3 – share of the five largest life insurance companies,

*B*4 – share of the five largest non-life insurance companies,

*B*5 – share of life premium in total premium,

[1] Further factors you can find in (Daly 2003).

[2] For computation we used figures from the year 1999 published in *European Insurance in Figures* (2001).

*B*6 – share of life investment in total insurance investments,
*B*7 – life investments/life premium,
*B*8 – non-life investments/non-life premium,
*B*9 – share of motor premium in total premium.

Variables eventually chosen to the study should have following features:[3]

- high inconstancy – we should omit all variables with low deviation ratio,[4]
- independency – it is necessary to eliminate those variables that repeat information carried by others.

The most popular correlation measure is Pearson correlation index. It detects direction and force of linear correlation between pairs of variables (Table 1).

Table 1. Correlation matrix

	*A*1	*A*2	*A*3	*A*4	*A*5	*A*6	*B*1	*B*2	*B*3	*B*4	*B*5	*B*6	*B*7	*B*8	*B*9
*A*1	1														
*A*2	0.72	1													
*A*3	0.13	0.48	1												
*A*4	0.67	0.56	0.52	1											
*A*5	0.99	0.69	0.16	0.63	1										
*A*6	0.68	0.72	0.60	0.95	0.65	1									
*B*1	–0.13	0.53	0.72	0.06	–0.11	0.23	1								
*B*2	0.91	0.65	0.22	0.51	0.95	0.54	–0.01	1							
*B*3	–0.17	–0.01	–0.11	–0.19	–0.18	–0.15	0.03	–0.10	1						
*B*4	–0.54	–0.04	0.27	–0.31	–0.55	–0.22	0.51	–0.45	0.73	1					
*B*5	0.70	0.76	0.10	0.31	0.66	0.38	0.14	0.63	0.30	0.07	1				
*B*6	0.53	0.70	0.26	0.20	0.48	0.34	0.26	0.50	0.06	–0.03	0.69	1			
*B*7	–0.23	0.46	0.64	–0.10	–0.18	0.13	0.98	–0.06	0.13	0.51	0.05	0.29	1		
*B*8	–0.12	0.33	0.10	–0.14	–0.08	–0.04	0.51	–0.02	0.53	0.61	0.36	–0.17	0.44	1	
*B*9	–0.37	–0.68	–0.46	–0.57	–0.35	–0.71	–0.56	–0.29	0.05	–0.02	–0.18	–0.16	–0.54	–0.31	1

From each highly correlated pair of variables $(r \geqslant 0.77)$[5] we have removed one variable – most correlated with all of the others variables. In one case we have not seen logical connection between correlated variables, so we decided to pay more attention to correlation diagram between them (*A*1 and *B*2 – premium/GNP and number of companies per thousand

[3] It is suggested in literature to leave in a study as small number of variables as possible. They should be representative for rejected ones, that means highly correlated with them. However, when the aim of a study is comprehensive analysis of insurance markets integration, it is useful to leave in a study as many variables as possible. The only condition is, that they care different information and are statistical correct (cf. Hellwig 1981, Nowak 1990).

[4] Kukuła (2000) suggests that deviation ratio should be over a dozen p.c.

[5] Correlation index lower than 0.77 means that more than 50% of information that variables take is unique.

inhabitants). We noticed one outlier in large distance from other observations.[6] When we omit the observation for Luxemburg, the correlation index was equal 0.19 (cf. Figure 1) – that means there is no necessity to loose any of those variables.

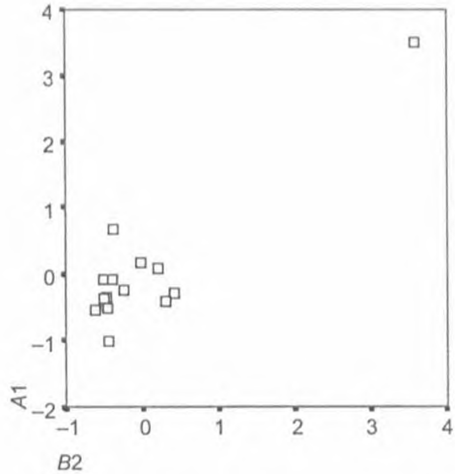

Fig. 1. Correlation between variables *A*1 and *B*2

Luxemburg has more outliers, and because most of the statistical methods are not resistant to outliers we have continued further analysis on fourteen countries without Luxemburg.[7]

Pearson correlation index do not detect non-linear correlation. Such dependency exists between variables when one of them is a negative and the second is a neutral.[8] The example of such dependency is presented in Figure 2. Correlation index equals 0.24, it means the independence of the variables. However curvilinear correlation index for function $y = |x-7.16| + 19.94$ equals 0.95 and for quadratic function 0.93. The analysis of correlation plots for all the pairs of our variables has not detected non-linear correlation.

---

[6] If there is an outlier in set of observations correlation index (for example) will be the higher the larger distance is.

[7] After omitting Luxemburg we have counted average value and standard deviation and standardized values one more time.

[8] A positive variable is one that high values remark high level of object's development, the opposite of a positive variable is a negative variable, where low values suggest low level of development. Neutral variable is a variable, that values from a definite range most appropriate and all the values higher or lower it describe lower level of development (cf. Michalski 2001).

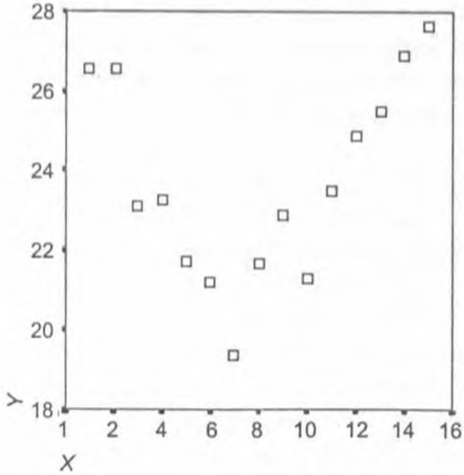

Fig. 2. An example of dependency between a negative variable (X) and a neutral (Y)

To confirm the choice among variables that we did, we have applied four multivariate statistical methods. First one is cluster analysis. In this method variables are assembled into larger and larger clusters using distance measure. So, the more similar variables mean the closer connection. To point out the non typical differences between objects in a multi-dimensional space we computed squares of Euclidean distances. Results of applying cluster analysis are presented in Figure 3. For example, the most similar are $A1$ and $A5$, so we should remove one of them. The results are very similar to those obtained in correlation matrix study.

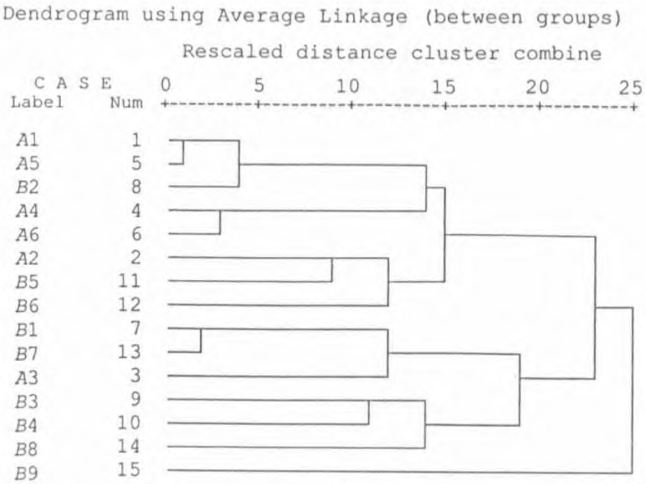

Fig. 3. Dendrogram using Average Linkage

Cluster analysis allows obtaining results easier and much faster but also does not detect curvilinear correlation and does not point out which of the variable reduce. As the result of cluster analysis we have also obtained distance matrix. The criterion of variables' reduction could be smaller sum of distances.

Second of applied clustering methods has been *k*-mean clustering. It divides variables into homogenic cluster, but you have to assume the number of them. Davies-Boulding index or comparison of variance inside the clusters and between clusters is helpful in this (Davis Boulding 1979). The most suitable number of clusters in our study was seven. Variables concerned in one cluster are similar to each other, so we can remove some of them. As the result of *k*-mean clustering we also obtained a table of distances from the center of the cluster. The less distance means the more variable is characteristic for the cluster. We can use it as the criterion of choosing the one, most characteristic variable if the cluster consists of more than two variables. Moreover the distance from a variable to a center of cluster points out the level of similarity. For example, the cluster number 4 is much more homogeneous than the cluster 6 (cf. Table 2).

Table 2. Results of *k*-means grouping

Variable	*Cluster*	*Distance*
*B*3	1	0.000
*A*2	2	1.564
*B*5	2	1.627
*B*6	2	1.759
*A*3	3	1.980
*B*1	3	1.033
*B*7	3	1.224
*A*4	4	0.582
*A*6	4	0.582
*B*9	5	0.000
*B*4	6	1.685
*B*8	6	1.685
*A*1	7	0.740
*A*5	7	0.331
*B*2	7	0.941

We can also observe distances between clusters. For example cluster 7 and 2 are quite close, so variables in those clusters are similar as well (cf. Table 3). Most distinct clusters are those indicated as four and five.

Table 3. Distances between clusters

Cluster	1	2	3	4	5	6	7
1		4.928	5.259	5.888	5.333	2.876	5.830
2	4.928		3.938	3.705	6.212	4.618	2.829
3	5.259	3.938		4.341	6.576	3.606	5.210
4	5.888	3.705	4.341		6.986	5.594	3.280
5	5.333	6.212	6.576	6.986		5.574	6.296
6	2.876	4.618	3.606	5.594	5.574		5.894
7	5.830	2.829	5.210	3.280	6.296	5.894	

Next method we used, was factor analysis. Its' aim is to express large number of variables by a few factors. The extraction of principal components amounts to a variance maximizing (varimax) rotation of the original variable space. First step of factor analysis is the scree test (cf. Figure 4). It is suggested to find the place where the smooth decrease of eigenvalues appears to level off to the right plot. According to this criterion we retained five factors.

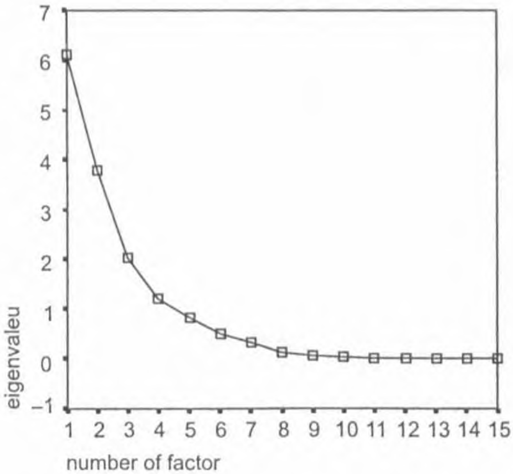

Fig. 4. The scree plot

We extracted five factors, they all accounted for 93% of variance. In Table 4 you can see principal components. Their analysis is next step of a study. Bolded figures are variables that are important parts of each factor, and they are similar to each other. So, we can reduce the number of variables, leaving one or two in each factor.

For example variables *A*1, *A*2, *A*5, *B*2, *B*5, *B*6 are collected in first factor what means that there are similarities between them.

Table 4. Principal components (Kaiser Varimax normalized)

Value	Principal component				
	1	2	3	4	5
A1	**0.880**	–0.173	0.387	–0.199	–0.011
A2	**0.762**	0.479	0.368	0.033	–0.082
A3	0.009	**0.652**	**0.607**	0.030	0.234
A4	0.331	–0.055	**0.918**	–0.109	0.032
A5	**0.874**	–0.145	0.359	–0.228	–0.051
A6	0.401	0.150	**0.885**	–0.083	0.021
B1	0.010	**0.966**	0.122	0.112	–0.104
B2	**0.862**	–0.044	0.251	–0.177	–0.022
B3	0.038	–0.061	–0.078	**0.948**	–0.058
B4	–0.312	0.433	–0.090	**0.825**	0.000
B5	**0.885**	0.075	0.036	0.369	0.032
B6	**0.696**	0.319	–0.045	0.074	**0.593**
B7	–0.028	**0.960**	–0.007	0.136	–0.021
B8	0.122	0.373	–0.099	0.576	**–0.667**
B9	–0.222	–0.470	**–0.615**	0.072	0.372

Quite different kinds of variables' grouping methods are neural networks. The one type of neural network could be applied is Self Organizing Map (SOM) (Kohonen 1997).

The great advantage of that method is much less sensitivity on data missing, outliers or non-linear correlation. However, the most difficult is the necessity of careful, empirical parameters choosing. It demands a large number of computations for different variants. The results obtained in that method are parallel to those obtained in *k*-mean clustering.[9]

[9] Results obtained by putting neuron network on the objects are finally grouped by *k*-mean clustering method. (number of cluster is computed by Davis-Boulding index).

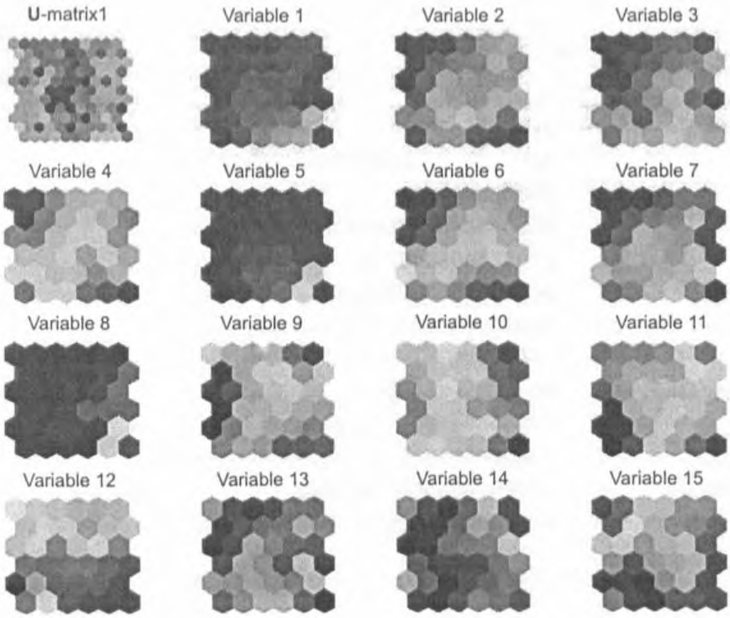

Fig. 5. SOM Diagrams

One of the results of SOM application is a neuron activity diagram. Diagram presents level of activation for each neuron, rich colors present higher activation. If we find two or more diagrams look the same it will mean that those variables are similar, for example *A*1, *A*5 and *B*2 the second *A*4 and *A*6, and the last one *B*1 and *B*7.

All applied methods pointed out that we should reduce from the study variables: *A*4, *A*5, *B*1.

3. STUDY RESULTS

To divide countries into homogeneous groups we applied cluster analysis. To point out existing differences between countries we used square of Euclidian distance. All the grouping methods gave similar results so the clustering is stable and reliable. As examples we present results of grouping using average linkage and Ward method (Figures 6 and 7).

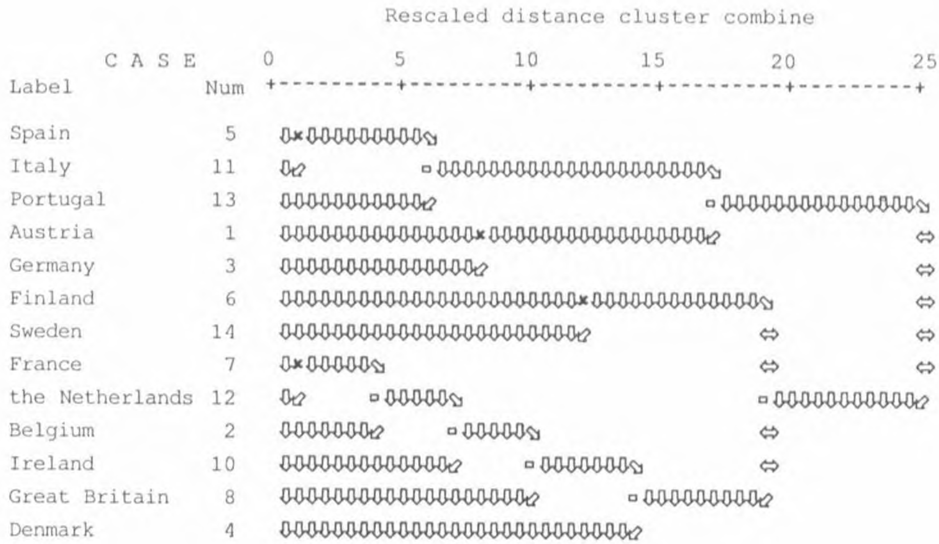

Fig. 6. Dendrogram using average linkage

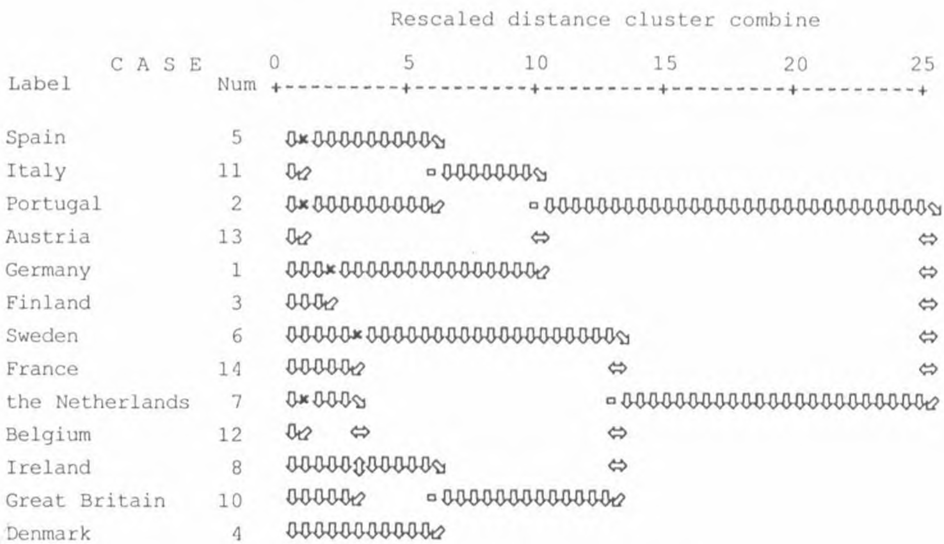

Fig. 7. Dendrogram using average linkage

In place of assuming number of clusters in *k*-mean method we focused on changes between clusters when the number of them have changed (cf. Figure 8). Extracting more than five clusters causes only a division of existing clusters, so five clusters seems to be optimal number.

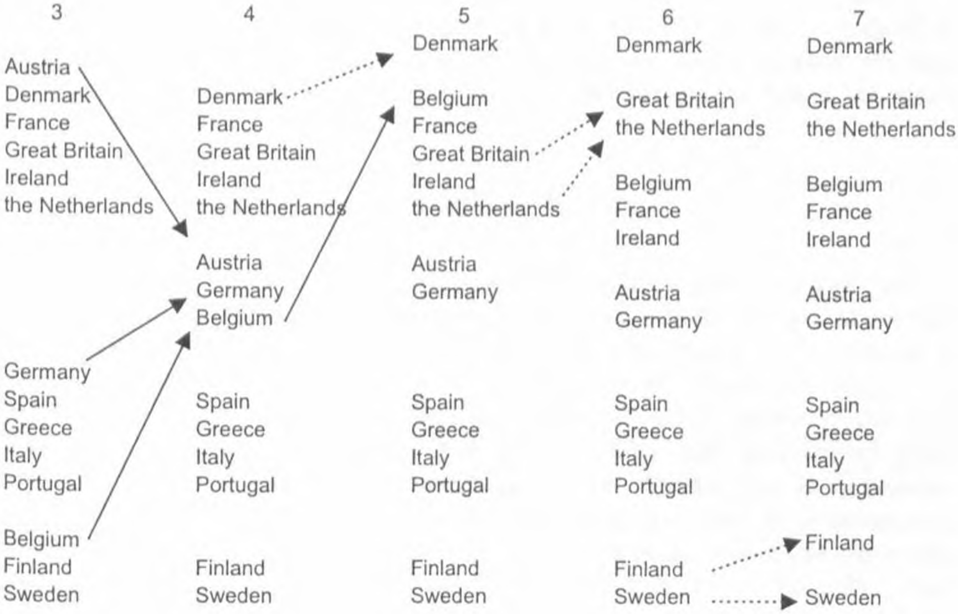

Fig. 8. Diagram of shifts between clusters

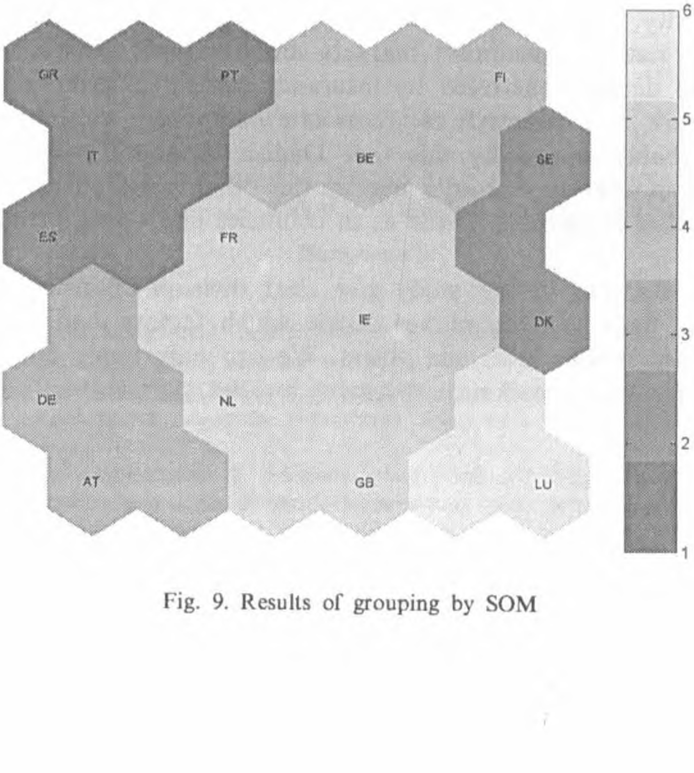

Fig. 9. Results of grouping by SOM

Finally, we used neural networks (cf. Figure 9). We noticed some differences from results obtained by previous methods. Most important is that Belgium and France is combined in one cluster, when other methods show large differences between them.

4. CONCLUSIONS

European insurance market is integrated on very low level.[10] We can recognize on it about five more homogenous clusters. Luxemburg is a country that made most by integration because of some reasons it is a country favorable for companies to establish their headquarters (Jurkiewicz, Wycinka 2003). Extremely high levels of all indicators for Luxemburg causes that this country could not be directly compared with other countries. If we consider remaining fourteen countries, a separate cluster is formed with Germany and Austria. The reason could be that in those countries insurance markets are highly developed, especially non-life markets. The second cluster is formed by Great Britain, Ireland, France, Belgium the Netherlands. In those countries rather life insurance markets pay an important role, especially as the retire products. Finland and Sweden have more restrictive finance law and this could be the reason of their similarity. Last group: Spain, Italy, Portuguese and Greece are countries where national insurance markets are developed least of all, those markets are deeply penetrated by insurance companies from other countries. Denmark is a country that form one component cluster. This situation is probably caused by this that Danish economy is comparable to those of Scandinavian countries and in the other hand an insurance market pay in this country such role as in countries like Great Britain, Ireland and so on.

Results obtained in this study give clear division of insurance market, however we have to by ourselves decide which factors made some of the markets more similar from the others. We can only apply our knowledge about the insurance markets.

[10] In the study "Dziesięć lat rozwoju jednolitego rynku ubezpieczeń" that has been broadcasted on the conference *Rynek usług ubezpieczeniowych w Unii Europejskiej – szanse i wyzwania dla Polski*, Gdańsk May 30th, 2003, we have compared the changes in rates describing insurance markets in 1990s. We concluded that the differences have not dwindle.

REFERENCES

Daly M. (2003), "Ireland: A Center for EU Life", *The Actuary*, 2.

Davis D. L., Boulding D. W. (1979), "A Cluster Separation Measure", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2, 224–227.

European Insurance in Figures (2001), Basic Data 2000/Complete 1999 Data, CEA.

Hellwig Z. (1981), *Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa.

Jajuga K. (1993), *Statystyczna analiza wielowymiarowa*, Państwowe Wydawnictwo Naukowe, Warszawa.

Jurkiewicz T., Wycinka E. (2003), "Analiza podobieństw rynków ubezpieczeniowych w krajach Unii Europejskiej i krajach kandydujących z wykorzystaniem metod analizy wielowymiarowej". In: *Inwestycje finansowe i ubezpieczenia – tendencje światowe a polski rynek, Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, **991**, 209–218.

Kohonen T. (1997), *Self-Organizing Maps*, Springer Verlag, Berlin–New York.

Kruczalak K. (1997), "Dostosowanie polskiego prawa ubezpieczeniowego do prawa europejskiego (prawa wspólnotowego)". In: Wąsiewicz A. (ed.), *Ubezpieczenia w gospodarce rynkowej*, cz. 3, Wydawnictwo Brauta, Bydgoszcz.

Kukuła K. (2000), *Metoda unitaryzacji zerowanej*, Wydawnictwo Naukowe PWN, Warszawa.

Michalski T. (2001), *Ubezpieczenia gospodarcze w Polsce i w Unii Europejskiej*, Difin, Warszawa.

Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa.

Świtał R. (1997), "Wspólnotowe prawo ubezpieczeń gospodarczych – ewolucja unormowań dyrektywalnych, orzecznictwo Trybunału Sprawiedliwości", *Wiadomości Ubezpieczeniowe*, **11/12**.

Tomasz Jurkiewicz, Ewa Wycinka

PROPOZYCJA WYKORZYSTANIA METOD ANALIZY WIELOWYMIAROWEJ W BADANIU STOPNIA INTEGRACJI RYNKÓW UBEZPIECZENIOWYCH

(Streszczenie)

Jednolity rynek ubezpieczeń funkcjonuje we Wspólnocie Europejskiej od 1993 r. (od wprowadzenia dyrektywy trzeciej generacji w ubezpieczeniach na życie). Jego działanie oparte jest na trzech podstawowych zasadach tworzących Unię Europejską, swobody tworzenia podmiotów gospodarczych, swobody świadczenia usług oraz przepływu kapitału między krajami.

Celem rynku wewnętrznego w dziedzinie ubezpieczeń jest zagwarantowanie wszystkim mieszkańcom Wspólnoty dostępu do możliwie najszerszej gamy wysokiej jakości produktów ubezpieczeniowych oferowanych przez zakłady ubezpieczeń z obszaru całej Wspólnoty. Ubezpieczyciele upoważnieni do działania w jakimkolwiek państwie członkowskim mogą prowadzić swoją działalność na terenie całej Wspólnoty i podlegają takim samym zasadom nadzoru. Gwarantowane dyrektywami jednakowe warunki rozwoju sektora ubezpieczeń nie znajdują pełnego odzwierciedlenia w regulacjach wewnętrznych państw członkowskich. Część z nich, poprzez sprzyjające regulacje podatkowe i administracyjne, jest zdecydowanie częściej wybierana

przez zakłady ubezpieczeń jako państwo siedziby. Pozostaje więc otwarte pytanie, czy europejski rynek ubezpieczeń jest organizmem jednolitym? Próba odpowiedzi na to pytanie wiąże się z oceną stopnia integracji rynków ubezpieczeniowych w Unii Europejskiej. Ocena taka nie może ograniczać się do analizy tylko jednego wskaźnika ekonomiczno-ubezpieczeniowego, gdyż zagadnienie integracji rynków ubezpieczeniowych, jak wskazano, jest zjawiskiem wieloaspektowym. W celu dokonania poprawnej oceny stopnia integracji rynków można posłużyć się metodami statystyki wielowymiarowej.

Celem niniejszego artykułu jest określenie stopnia integracji rynków ubezpieczeniowych krajów Unii Europejskiej przy wykorzystaniu metod analizy wielowymiarowej. Równoległym celem jest porównanie wyników uzyskiwanych przy pomocy różnych metod statystyki wielowymiarowej i próba ich oceny.