

# Methodological approaches to identification of clusters in regional economy system

Angela Mottaeva<sup>1,\*</sup>

<sup>1</sup>Moscow State University of Civil Engineering, Yaroslavskoeshosse, 26, Moscow, 129337, Russia

**Abstract.** The cluster (cluster group) identification methodology according to Porter includes some successive steps. The basis of the economic agglomeration identification methodology is the method of distribution of employment over industries on the territory, according to which: first, all sectors (industries) are grouped into three types; second, the industries are identified; third, the cluster group composition is determined; fourth, identification of the most significant cluster groups, precluding false correlations between industries, is performed; fifth, the existing intersections in cluster groups are analysed. The authors present the algorithm of identifying the developed region clusters is based on identification of agglomeration effects of concentration, urbanisation and joint localisation of regional industries' enterprises and their economic efficiency, thus, the algorithm is the scientifically grounded method of forming the regional cluster structure. The use of this algorithm will enable developing the efficient cluster policy aimed at increase of employment in the region, salary, wage and budget income level, growth of the competitive ability of the region.

## 1 Introduction

The sources of the cluster approach were founded in the works of English economist Alfred Marshall. In Principles of economics [1-7] he introduced the concept of an industrial district – a pattern of organisation of manufacture based on geographical concentration of companies of the same industry applied in industrial cities of England.

The concept of cluster was introduced into the economic theory by the American economic Michael Porter [11, 12]. The cluster approach is widely spread in foreign countries [5]. The development of cluster policy in developed countries' economies was facilitated by some regulatory documents accepted in the second half of 2000's (Manifest of clusterisation in Eurozonecountries (2006), European Cluster Memorandum). The main task of clusterisation, as prescribed in the above mentioned documents, is to increase the quantity of clusters and cluster groups of enterprises for the purpose of enhancing

---

\* Corresponding author: [angela-1309.m@yandex.ru](mailto:angela-1309.m@yandex.ru)

competitive ability of national economies. To transfer to the innovative type of economic development, improvement of the population life quality based on assuring of the new quality of economical growth, forming stable prerequisites for the global competitive ability of Russia, by Decree of the Government of Russian Federation dated the 17th of November, 2008 the "Concept 2020" was approved, the concept of social and economic development of Russia till 2020. This is the first regulatory legal document which notes the significance of formation of territorial production clusters in priority branches of economy, and the framework of cluster policy in Russian Federation is established.

## **2 Materials and Methods**

Cluster (cluster group) identification methodology by M. Porter includes some successive steps which are characterised below according to [11, 19].

The methodology enables assessing the strength and significance of all cluster groups existing on the particular territory (region, macro-region) and performing only approximate identification of cluster groups which is sufficient for provision of efficiency of regional cluster policy. The calculation of the significant cluster groups is possible based on the statistical data on different directions enabling more precise assessment of the cluster potential. Using this index, it is possible not only to assess the development of the cluster group, but also to determine the development proportions.

The first step is division of the territory in question into regions. Individual states of the USA are considered as such a territory in the work by M. Porter. In accordance with the research objective, any territories may be treated as regions: states, counties, lands, etc.

The basis of the economic agglomeration identification is the method of distribution of employment over industries on the territory.

The first step – all sectors (industries) are grouped into three types:

- Branches of the local economy in which employment is proportional to population size. The main feature of the local economy branches is their focus on the internal (local) demand of population and enterprises (health care, household service, trade and public catering, education, etc.);
- Extractive and processing industries where employment is concentrated mainly in locations of mineral resources (oil, gas, coal, wood). Such industries are characterised by competition both on the internal and external market;
- The basic industries where employment is associated with multiple factors which determine competitive ability of the territory (machine-building, automotive assembly, etc.). Such industries are associated with international and interregional trade.

The second step includes identification of industries (sectors). According to the majority of researchers [3, 4, 10, 12], identification of sectors of branch directions is not very complex. However, identification of basic industries and local economy branches in practical activity requires elaboration of some additional assessment criteria [4, 20].

The basic economy branches have the following three characteristics:

1. The average value of localisation coefficient (LQ) including five regions having maximum values of this criterion within boundaries of each of the industries under consideration. The national economy branches where the identified average value of the localisation coefficient is two or more are tradable (resource-dependent):

$$LQ = \frac{Emp_{ig}}{Emp_g} / \frac{Emp_i}{Emp} \quad (1)$$

Where  $LQ$  is "Localisation coefficient";  $Emp_{ig}$  is the employment volume in cluster group  $i$  in region  $g$ ;  $Emp_g$  is the total employment volume in region  $g$ ;  $Emp_i$  is the employment volume in cluster group  $i$ ;  $Emp$  is the total employment volume.

2. Gini coefficient which is calculated for each industry individually. Industries with 0.3 and higher values of this coefficient are considered to be resource-dependent ones, as such a non-uniform distribution means the industry under consideration is highly concentrated in several regions:

$$G_i = \frac{1}{2} \frac{1}{\sum_r \frac{1}{R} \Pi_r} \sum_i \sum_s \frac{1}{R^2} \left| \frac{L_{ir}}{\Pi_r} - \frac{L_{is}}{\Pi_r} \right| \quad (2)$$

Where  $G_i$  is Gini economic inequality index;  $L_{ir}$  is the employment volume in cluster category  $i$  in region  $r$ ;  $\Pi_r$  is the employment volume in all cluster groups in region  $r$ ;  $R$  is the quantity of regions in the country.

3. Determination of  $LQ$  index – the share of all employment volume in all regions with coefficient value "greater than or equal to 1" in employment rate at the national economy level. Calculation of this criterion requires separation of all territories with localisation coefficient for particular industry, which is equal or greater than one. Then all indices of employment in the industry under consideration on the particular territories (regions) are summarised. If this criterion is at least 50% of the national employment rate in the same industry, it shall be referred to resource-dependent industries.

Such approach is necessary for identifying resource-dependent industries and industries which depend on natural resources, separating them from industries of the local economy in which the employment rate is distributed uniformly. In accordance with the presented methodology, the enterprises' cluster groups include only resource-dependent (tradable) industries which are free to select their location [3,15,18].

The third step is associated by Harvard school representatives with solution of the task within which the cluster group composition is determined. That is to say, the tradable industries included into the cluster group are determined. Due to multilateral nature of relationships between companies of different industries of the regional economy, it is rather difficult to identify positive externals as a criterion of assignment to the same group. Usually, the repeatedly occurring state-to-state geographical proximity of companies referred to different tradable industries is selected to be the criterion. This approach is mainly based on the fact that the US economy is the largest in the world and also it includes the greatest amount of industries. It is also not insignificant that the spatial US economy is a single entity, and no impediments occur when companies are transferred from one state territory to another. It results in the possibility to identify stable combinations (patterns) of co-localised branches forming cluster groups.

The fourth step of methodology identifying the most significant cluster groups in the US economy is elimination of any false correlations between branches. False correlations are cases where joint localisation of branches is not the single cluster group. Elimination of false geographical correlations can be performed based on the analysis of the inter-industrial balance data and the use of case-study method for qualitative analysis of actually existing regional interactions in historical perspective.

The Harvard school representatives have obtained the following results: the quantity of the identified cluster groups is 41, including the average of 29 branches per group.

At the fifth step, the Harvard school representatives have conducted the analysis of the existing intersections in the cluster groups. If the cluster groups intersect, then a situation occurs where it may be stated that one industry is included into several cluster groups. Besides, some subgroups including industries with maximum correlation are identified within the cluster group.

An advantage of Harvard school representatives' approach is the possibility to carry out the cluster group activity analysis at once, excluding analysis of the industry activity types. In this case the groups of interconnected activity types combined into patterns of joint industry localisation were considered as cluster groups.

Identification of industrial clusters remains to be the subject of scientific research works. In particular, methodology of M. Porter was creatively used by European Cluster Observatory [2]. This organisation carried out mapping of economic agglomerations of European countries included into EU. The modification of Porter's approach was based on comparison of industry classifier of the USASIC and European classifier NACE. After identification of economic agglomerations they were analysed using indices "Localisation coefficient" (3), "Size" (4), "Focus" (5).

$$LQ = \frac{Emp_{ig}}{Emp_g} / \frac{Emp_i}{Emp} \quad (3)$$

Where  $LQ$  is localisation coefficient;  $Emp_{ig}$  is employment volume at enterprises of cluster group  $i$  in region  $g$ ;  $Emp_g$  is the total employed population volume in region  $g$ ;  $Emp_i$  is the employment volume in  $i^{th}$  cluster group;  $Emp$  is the total employment volume in the national economy.

$$Size = \frac{Emp_{ig}}{Emp_i} \quad (4)$$

Where  $Size$  is the magnitude of  $i^{th}$  cluster group.

The other system indices are described above.

$$Focus = \frac{Emp_{ig}}{Emp_g} \quad (5)$$

Where  $Focus$  is the centre of cluster group concentration (focus).

The other system indices are described above.

According to G. Lindqvist (European Cluster Observatory) [4], the threshold values for significant groups of the region are as follows:

1 – value of index  $3 \geq 2$ ;

2 – indices 4 are of leading character (the regions must be referred to 10% leader regions in terms of this index);

3 – indices 5 are of leading character (the regions must be referred to 10% leader regions in terms of this index).

If this index matches to the criterion, then "a star" is assigned to the cluster group. The maximum of three stars may be assigned. The greater is the quantity of "stars", the higher is the group strength.

### 3 Results

Thus, the methodology allows assessing the strength and significance of all cluster groups existing on the particular territory (region, macro-region).

However, this methodology has got some drawbacks, since:

- it is unable to assess mutual priority of the cluster groups;
- it does not provide tools for assessing the efficiency of activity of the enterprises included into the cluster groups;
- it does not assess the tightness of interrelationships between the cluster groups;
- it is based on limited range of indices (three), it takes into consideration rigid quantitative limitations and prefers employment indices.

Consequently, this methodology only allows carrying out approximate identification of cluster groups which is not sufficient for assuring efficiency of regional cluster policy.

Let us refer to the research works conducted by Ye. S. Kutsenko in cooperation with a group of Russian scientists [1, 17]. It suggests a comprehensive aggregative index of the cluster group development (Table 1) enabling a more precise assessment of the cluster potential. The comprehensive index is based on a wider range of indices of the official statistics. Using this index, it is possible not only to assess the development of the cluster group, but also to determine the development proportions.

Indices of significance of a cluster group are calculated using formulas (3, 4, 5) based on statistical data and combine such indices as: employment volume; amount of shipped products (cost index).

Indices of significance of a cluster group are calculated using formulas (3, 4, 5) based on statistical data and combine such indices as: employment volume; amount of shipped products (cost index).

**Table 1.** An aggregative index of the cluster group development level: structure and contents.

| <b>Groups of indices</b>                          | <b>List of indices</b>   |
|---|--|
| Cluster group significance indices                | <ul style="list-style-type: none"> <li>• cluster group significance assessment (it combines assessment of localisation, size, focus calculated based on the employment indices and the cost of the shipped products);</li> <li>• cluster group uniqueness assessment.</li> </ul> |
| Cluster group coherence indices                   | <ul style="list-style-type: none"> <li>• the quantity of intersections of significant cluster groups;</li> <li>• indices of concentration and urbanisation of a cluster group.</li> </ul>  |
| Indices of economic efficiency of a cluster group | <ul style="list-style-type: none"> <li>• average monthly salary over the cluster group;</li> <li>• profit over the cluster group;</li> <li>• investment over the cluster group.</li> </ul>   |

The significant cluster group calculation is possible based on the statistical data on different directions. However, as it has already been shown, the employment statistics (the most stable index) is the priority. Along with this, to some extent it is interesting to determine the significant cluster groups by the shipped product (carried out works, provided services) amount index. Such index reflects those cluster groups in which the city is leading in terms of the shipped products (carried out works, provided services). It is obvious that the groups significant in terms of employment volume and shipment amount should at large correspond to each other (the greater employment level results in a greater shipment

amount), but some exceptions are possible. Within this research, in accordance with methodology of M. Porter and European Cluster Observatory [2,12,16], the employment statistics adjusted based on the results received from the research into the shipment indices was assumed as the basis.

Index of uniqueness of a cluster group reflects the degree of cluster group employment non-uniformity by regions.

The greater is the value of non-uniformity of the cluster employment volume distribution, the more significant is its position in the region. This index is correlatable and is calculated in points in the interval from 0 to 1. To calculate the aggregative index of the uniqueness, formulas (6), (7), (8) are used. They reflect the employment volume distribution non-uniformity degree in cluster groups.

Gini coefficient must be used as an inequality level index. Gini coefficient is a statistical index of the degree of stratification of some aggregate in relation to some attribute in question[6]. In this case the calculation formula looks as follows:

$$G_i = \frac{1}{2} \frac{1}{\sum_r \frac{1}{R} \frac{L_{ir}}{\Pi_r}} \sum_r \sum_s \frac{1}{R^2} \left| \frac{L_{ir}}{\Pi_r} - \frac{L_{is}}{\Pi_r} \right| \quad (6)$$

Where  $G_i$  is Gini economic inequality index;  $L_{ir}$  is the employment volume in cluster category  $i$  in region  $r$ ;  $\Pi_r$  is the employment volume in all cluster groups in region;  $R$  is the quantity of regions in the country.

Using Gini coefficient, it is possible to assess the inequality between the upper and lower (in terms of income) population groups, but it is impossible to determine the location where the inequality originates on Lorenz curve. Theil index [6] can be used for this purpose which is calculated by the following formula in this case:

$$T_i = \sum_r \frac{1}{R} \frac{\frac{L_{ir}}{\Pi_r}}{\sum_r \frac{1}{R} \frac{L_{ir}}{\Pi_r}} \ln \left( \frac{\frac{L_{ir}}{\Pi_r}}{\sum_r \frac{1}{R} \frac{L_{ir}}{\Pi_r}} \right) \quad (7)$$

Where  $T_i$  is Henri Theil economic inequality index;  $L_{ir}$  is the employment volume in cluster category  $i$  in region  $r$ ;  $\Pi_r$  is the employment volume in all cluster groups in region;  $R$  is the quantity of regions in the country.

Herfindahl index is generally used for assessing the degree of monopolisation of the industry, i.e. the index reacts to the market share of each company in the industry [8]:

$$H_i = \sum_r \left( \frac{L_{ir}}{\sum_r L_{ir}} \right)^2 \quad (8)$$

Where  $H_i$  is Herfindahl economic inequality index;  $L_{ir}$  is the employment volume in cluster category  $i$  in region  $r$ .

The analysis of the index of intersection of significant cluster groups with other cluster groups is based on the cluster group intersections "map" created by M. Porter [12]. Intersections mean activity types shared by the cluster groups. The greater is the amount of intersections between the significant cluster groups, the better it is, since in such a case the development of one cluster group will be accompanied by the development of other groups associated with it. The cluster group coherence data can enhance the cluster policy efficiency due to focusing efforts on the intersecting cluster groups. The design formulas of concentration (9) and urbanisation (10) indices are provided below.

$$Conc = \frac{Q_i}{Q_{ix}} \quad (9)$$

Where Conc is the cluster group concentration level;

$Q_i$  is the ratio of the quantity of links between employment volume in the cluster group within the region to the quantity of links between employment volume in that cluster group in the country;

$Q_{ix}$  is the ratio of the quantity of links between employment volume in the cluster group with employment volume in other cluster groups within the region to the quantity of links between employment volume in that cluster group with employment volume in other cluster groups in the country.

$$Urb = \frac{Q_{ix}}{Q_x} \quad (10)$$

Where Urb is the urbanisation level of the cluster group;

$Q_{ix}$  is the ratio of the quantity of links between employment volume in the cluster group with employment volume in other cluster groups within the region to the quantity of links between the employment level in that cluster group with the employment level in other cluster groups in the country;

$Q_x$  is the ratio of the quantity of links between the total employment volume within the region to the quantity of links between the total employment volume in the country.

The localisation potential is the product of these indices:

$$Loc = Conc \times Urb \quad (11)$$

The cluster group localisation potential index shows the degree of attraction of those employed in a cluster group to joint localisation with each other (the cluster group concentration level), and the degree of attraction to regions with high employment volume in general (cluster group urbanisation level). It is obvious that the higher cluster group concentration and urbanisation indices reflect the presence of a significant positive effect of the joint localisation. Such a cluster group has high potential of inner coherence and integration into the region economy. It was supposed that one link is formed by two employed persons inside one definite category or between two definite categories (region, cluster group in a definite region, the whole country, etc.). For example, if ten people are employed in the category under consideration, then each of them will form a link with remaining nine people, and the total amount of links will be  $10 \times 9 = 90$ . If two categories in each of which ten persons are employed are considered, then each person from the first category will form a link with each person from the second category, and the total amount of links will be  $10 \times 10 = 100$  [4,12,13]. The economic efficiency indices are based on the comparative analysis of the level of salary, investments and profit in each cluster group to range them and to determine the clusters ensuring the maximum contribution to the region economy development. The source information base for the index analysis is statistical information of the Federal State Statistics Service (Rosstat) based on activity types of All-Russian Classifier of Types of Economic Activity on the following indices: "Average monthly accrued salary per one employee over the whole range of organizations"; "Long-term investment for the ordinary activity types"; "Profit (loss) on sales based on accounting data over the whole range of organizations" indexed taking into consideration the index "manufacturers' price indices per economic activity type over the period from the beginning of the reporting year as % of the respective period of the previous year".

## 4 Discussion

Taking into consideration the significant amount of works on the issue of identification of regional cluster groups, the problem of forming the innovative type economy by forming regional clusters remains to be topical [9]. The problem of cluster identifications remains to be the key issue in the cluster theory. In accordance with foreign experience, the methodological tools used for it vary broadly in practice. The foreign experience enables identifying paired categories to characterise processes and technique of identification of enterprise cluster groups. Alongside with that, many of them are based on two main approaches. The first of them (often termed "approach from below") suggests that it is only necessary to identify clusters in a definite region (particular territory) in which some dominating branches (leader industries) and particular enterprises are present. The second approach uses the methodology ("approach from above") based on search and identification of special organization of the product manufacture (service provision) which is focused on specific types of economic activity. Then a nation-wide specimen is used for identification and primary analysis of particular regional clusters.

The majority of specialists hold to the opinion that it is more appropriate to identify clusters "from above" by using synthetic approach, that is combination of two approaches – spatial and industrial. The Harvard school representatives, and primarily M. Porter, use this synthetic approach which has been recognized as the classical one and is widely used by specialists of other countries [11, 14, 20].

## 5 Conclusions

The analysis of the existing approaches to identification of regional industrial clusters has shown that in conditions of limited information base, i.e. absence of statistical information about inter-industrial balance indices in Russian Federation for determining the cluster structure of the region, i.e. identification of key industrial and inter-industrial groupings of the region, it is possible to use only the approach identifying "reference" clusters. The "reference" clusters are the most probable agglomerations of related activity types. The composition and structure of "reference" clusters has been developed in European and American economy based on the analogues of All-Russian Classifier of Types of Economic Activity – industrial classifier SIC (USA) and classifier NACE (Europe). The use of this approach is possible by adapting the structure of "reference" cores of clusters for Russian economy in accordance with All-Russian Classifier of Types of Economic Activity which allows determining the composition of the most probable cores of cluster groups.

To assess the region cluster development level, it is suggested to use criteria of European Cluster Observatory which enables assessing the cluster groups strength based on the employment statistics and identify the regions in which the clusters under consideration are significant. Identification of such regions allows drawing a conclusion on the existence of competing centres of concentration of the respective industries – cluster groups capable of increasing the employment rate at the expense of other industries as a result of action of positive agglomeration externals on the territory under consideration. It is suggested to expand the The European Cluster Observatory's methodology by analysing the following indices:

- Indices of uniqueness which reflect the cluster group employment non-uniformity degree by regions;



- Coherence indices which reflect the degree of attraction of people employed in the cluster group to joint localisation and the degree of attraction to regions with high employment volume;
- Economic efficiency indices, including level of salary, investments and profit allowing to determine the clusters ensuring the maximum contribution to the region economy development.

The developed algorithm of identifying the developed region clusters is based on identification of agglomeration effects of concentration, urbanisation and joint localisation of regional industries' enterprises and their economic efficiency, thus, the algorithm is the scientifically grounded method of forming the regional cluster structure. The use of this algorithm will enable developing the efficient cluster policy aimed at increase of employment in the region, salary, wage and budget income level, growth of the competitive ability of the region.

## References

1. T.P.Danko, E.S.Kutsenko, "ROST" **1**, 248-252 (2012)
2. European Cluster observatory (2012)
3. D.F. Heany, G. Weiss, Journal of Business Strategy, **4 (1)** (1983)
4. G. Lindqvist, *Disentangling Clusters Agglomeration and Proximity Effects* (2009)
5. E.B. Lenchuk, The editorial Board of "Federalism", **5** (2012)
6. M.E. Porter, Regional Studies **4**, 549-578 (2003)
7. A. Marshall, *Principles of Economics* (Macmillan and Co. Ltd., London, 2008)
8. L.S. Markov, V.M. Markova, Vestnik NSU, **12-1** (2012)
9. M. E. Porter, *On Competition* (Harvard Business School Press, Boston, 1998)
10. N.G. Verstina, E.M. Akimova, T.N. Kisel, Asian Social Science, **11-14** (2015)
11. N.G. Verstina, T.S. Meshcheryakova, Biosciences Biotechnology Research Asia, **12 - 2** (2015)
12. A.Mottaeva, MATEC **73**, 07020 (2016)
13. A.B. Mottaeva, A. Zheltenkov, I. Stukanova, S. Ryabichenko, S. Zhuk, MATEC **73**, 07026 (2016)
14. I. Polyakova, E. Vasilyeva, Procedia Engineering, **165** (2016)
15. L.V. Nikolova, D.G. Rodionov, A.B. Mottaeva, International Journal of Economics and Financial Issues **6(S3)**, 1-7 (2016)
16. An.B. Mottaeva, As.B. Mottaeva, International Journal of Applied Engineering Research **10- 23**, (2015)
17. An.B. Mottaeva, As.B. Mottaeva, International Journal of Applied Engineering Research, **11- 9** (2016)
18. E. Stein, N. Taskaeva, E. Chibisova, Procedia Engineering, **165** (2016)
19. T. Miroshnikova, N. Taskaeva, MATEC **73**, 07006 (2016)
20. A. Pimenova, S. Kuzmina, N. Morozova, A. Mottaeva, MATEC **73**, 07018 (2016)