CLUSTER POTENTIAL IDENTIFICATION TOOL SUPPORTING THE PROCESS OF RELATED VARIETY DISCOVERY

Maciej Piotrowski – Korneliusz Pylak – Michał Thlon

Abstract

Clusters are becoming an increasingly popular concept in the last years which is reflected in a growing number of clusters support policies and initiatives on EU, national and regional level. It is widely accepted, that clusters can contribute to the improvement of competitiveness and innovativeness of the regional economies. Recently, however, not only clusters are gaining importance, but mainly related variety, defined by Boschma as 'industrial sectors that are related in terms of shared or complementary competences' (2007). Related variety is a key concept of creating regional smart specialisations because of its positive impact on regional development. We claim that cluster potential identification tool can still be one of the most effective tools for identifying related variety and specialisation within it on the micro level. The literature review shows several methods, that are widely used to identify cluster potential, among them exist such as: case studies, concentration analysis and input-output methods. This papers presents new approach that may be used to find related sectors cooperating together not only between regions, but also within the region. The approach seeks for the real functioning cooperation links, not only existing sectors in given proximity. Finding these links will allow to identify related sectors that are exchanging knowledge through value chains. At the end of the paper authors present the verification of the model on the base of the agrifood industry in Poland.

Key words: clusters; related variety; potential identification; mathematical model; case study

JEL Code: O21, O31, R11

Introduction

Clusters are becoming an increasingly popular concept in the last years, what is reflected in a growing number of clusters' support policies and initiatives on EU, national and regional level (Asheim, Moodysson, & Tödtling, 2011). It is widely accepted, that clusters can contribute to the improvement of competitiveness and innovativeness of the regional economies (Asheim, Cooke, & Martin, 2006).

The paper discusses the ideas of prospective clusters identification, that are broadly considered as the way to increase the innovativeness and competitiveness of the regions. However in the period 2007-2014 a number of unsuccessful cluster initiatives could be observed. One reason for its failure was improper identification of its potential resulting in low level of theirs sustainability, typically resulting in the cease of activities shortly after external financing has been terminated (Rubach, 2013). The paper discuses new approach merging the best existing mathematical models for better identification of clusters in the future.

The paper is structured as follows. The next section explores and clarifies the theoretical foundations on cluster identification process and tools. The third section lays out the empirical design, including hypotheses, preparation of data and measurement. The fourth section presents and discusses the empirical results of analysis (based on the analysis of quantitative data on agrifood industry in Poland); and the fifth offers a conclusion.

1 Cluster identification process

In the literature, theoretical and conceptual issues of clusters identification seem to be incoherent (Fornahl, Henn, & Menzel, 2010), although this fuzziness can be seen as a part of creation phenomena, since creative process cannot be fitted in only one or even few paradigms. We agree that clusters have to be discovered, not built, so deep regularities in the economy structure have to be investigated (Bresnahan & Gambardella, 2004), including related variety, knowledge spillovers, absorptive capacity against the background of social capital, technology branching, spin-off subsidies and entire regional innovation system (Fornahl et al., 2010). For a local cluster to be created, sufficient prerequisites have to be given in a region, that are usually different in different industries and exist before cluster emerges (Brenner & Muehlig, 2013). These prerequisites can be traced ex-post, when some initial networks are already created and some missing links could be supported and encouraged (Crespo, Suire, & Vicente, 2014).

The literature review shows several methods, that are widely used to identify cluster potential, i.e. concentration analysis (Alonso-Villar & Del Río, 2011), input-output methods (Lehtonen & Tykkyläinen, 2014) and case studies (Brenner & Muehlig, 2013; Fornahl et al., 2010). Two first are quantitative methods are the latter is qualitative one. These qualitative methods, however, are not related directly to specific networks, but analyse concentration of companies or goods exchange. And what is worth noting, relational mix of the networks is crucial, not just the network size alone (Lechner, Dowling, & Welpe, 2006) and the network aims change through time, so is the strength of ties (Lechner & Dowling, 2003).

This statement leads us to the notion of related variety, defined by Boschma and Iammarino (2009) as '*industrial sectors that are related in terms of shared or complementary competences'*, focusing on knowledge networks more than on business networks. This approach indicate not only location and spillover externalities, but also social and knowledge values that can be measured (Ter Wal & Boschma, 2009; Walter, Lechner, & Kellermanns, 2007). Then, it is worth noting, that basing on an assumption similar knowledge pool is required to produce products or services that are mostly commonly sold outside the country, Hidalgo, Barabási, Winger, and Hausmann (2007) created their product space theory. This approach, fitted to individual clusters and to value chains (i.e. supplier – buyer relationships or company – scientists) takes into account all strengths of abovementioned measures and techniques.

2 Research design

2.1 General approach

This papers presents new approach that may be used to find related sectors cooperating together not only between regions, but also within the region. The approach seeks for the real functioning cooperation links, not only existing sectors in given proximity. Finding these links will allow to identify related sectors that are exchanging knowledge through value chains.

The approach discussed in this paper adopts triangulation of research methods combining both primary and secondary data sources. The process consists of the following stages:

1. Identification of the key industries within which the clusters could be established with the adoption of concentration analysis based on location quotient (LQ). This step widely used in practice is largely dependent on the data available for an analysed area with the use of NACE Rev.2 classification for a chosen variable (it can be e.g. employment, number of companies, industry added value). Typically for regions in the EU only data for level 2 of the Nomenclature of Territorial Units for Statistics is available, therefore results of this step are relatively general from both perspectives: industry classification and geographic location. The next step presents authors extension to the LQ models.

2. For each selected industry field, a representative sample of companies is surveyed with the use of computer assisted telephone interview (CATI) or similar method. Data collected should contain at minimum such variables as: industry represented and industry partners separately on the supply-side and demand side coded with minimum detail level equal to 4 digits of NACE classification. For more extensive analysis involving geographical proximity (not discussed in this paper due to its complexity), it is advised that each cooperation link will be coded as local (NUTS 3), regional (NUTS 2), national or international level.
3. Based on CATI results there can be created a measure inferring the relatedness between companies and their suppliers and clients referring to their 4 or 5-digit NACE code. To quantify this relatedness there has to be assumption given that if two sectors are often cooperating together as a supplier – buyer, their businesses are similar and can cooperate together with some level of probability. If the cooperating businesses are from different groups of sectors, their relatedness can be called related variety. What's more, in that case there will be direct indication for possible cluster sectors picking up and the diversity of cluster can be estimated. To make final products we need chunks of embedded knowledge which we call capabilities. We assume if companies are cooperating together, selling and buying semi-products from each other, they share capabilities needed to produce one good or the series of goods. If a given majority of companies from one sector cooperate with other sector it means they increase, complement or just use capabilities from the other sector. Since we do not observe capabilities directly, we create a measure that infers the similarity between the capabilities possessed by a pair of companies by looking at the probability that they are cooperating. To quantify this similarity we assume that if two companies need the partners' capabilities, they will cooperate in different level, as a supplier or buyer. By the same token, companies that are not related and which capabilities are different and not related are less likely to cooperate.
4. Calculated the set of measures inferring the relatedness between companies representing different, but related industries is the base for identification of the business value networks that are defined by high value of the measures between analysed companies. These identified value networks can be the foundation for establishing successful cluster initiatives, since they already organized strong cooperation links (based on the goods, service or revenue flows or other intangible

values). For better visualisation, this step can be supported by the graphing of the value network with the use of various techniques and tools.

2.2 Hypothesis

The proposed approach to identify clusters with adoption of related variety discovery can significantly contribute to the development of regional and national cluster policies. This can be assured through better identification of cluster potential by extending existing methods and thus assuring higher sustainability of supported cluster initiatives. Additionally the model can lead to the identification of industry gaps, that might hinder functioning of the cluster. Since the usefulness of the model outlined with the use of the hypothesis can not be statistically tested, authors decided to validate the correctness of the model with the use of regional data for Lubelskie region in Poland.

3 Empirical results

3.1 Data collection and analysis

Within the frames of the scientific cooperation with the Marshal Office in Lubelskie Region, authors attempted to identify cluster potential in chosen industries. Following the first step a number of industries with high concentration of companies were identified using location quotient analysis. The one selected for further analysis was *Crop and animal production* (division 01 according to NACE Rev.2). Next study of 2000 companies in the region representing this industry and related ones (e.g. Food processing) were analysed with the use of CATI surveying technique. Based on CATI results a set of measures inferring the relatedness between companies and their suppliers and clients referring to their 4 or 5-digit NACE code was calculated. The measure is based on the conditional probability that the sector s will cooperate with sector s'. Since conditional probabilities are not symmetric, the minimum of the probability of cooperating sector s with given sector s' and the reverse was taken, to make the measure symmetric and more stringent. For example, the company from sector 01.11.Z occurs 364 times, suppliers from sector 46.21.Z occurs 321 times, but these suppliers supplied companies from sector 01.11.Z only in 194 cases. Then, the proximity between these two sectors is equal to 194/364=0.53. Note that there is 194 divided by 364 instead of 321 to minimise false positives. Formally, for a pair of sectors and the proximity is defined as:

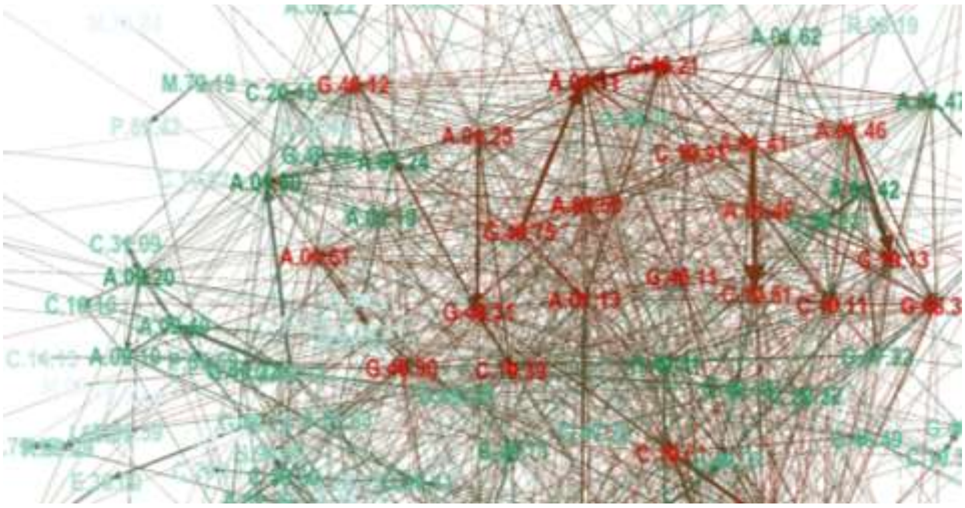
$$\phi_{s\,s'} = \frac{\sum_i M_{si} M_{s'i}}{max(c_s; c_{s'})}$$

where M_{si}=1 if company from sector s occurs in the case i and 0 otherwise; c_s is the complexity of sector s. High values of the measure (arbitrarily in this case as equal or higher to 0,7) illustrate strong cooperation links between given industries that can be selected for further analysis with the use of visualisation techniques.

3.2 Visualisation techniques and tools

Identified value networks, that can be transformed into successful cluster initiatives can be visualised using various techniques and tools The following figures illustrates a chunk of the generated overall diagram with the use of Gephi software. The legibility of the diagram is hindered, since for the analysis we took 2000 companies representing 200 cooperating industries (coded with 4-digits NACE Rev. 2). 20 existing value networks have been identified concentrated along such products lines like e.g. fruits, hop, piglets, tobacco production.

Figure 1. Sample overall diagram illustrating cooperation links between industries (4-digits NACE Rev. 2).



The most occurring links can be extracted to further process with the use of graphing techniques. The following figure is the illustration of the final results obtained, that can be used by the decisive bodies to support cluster initiative establishment.

Figure 2. Sample diagram illustrating value network in the area of fruit production.

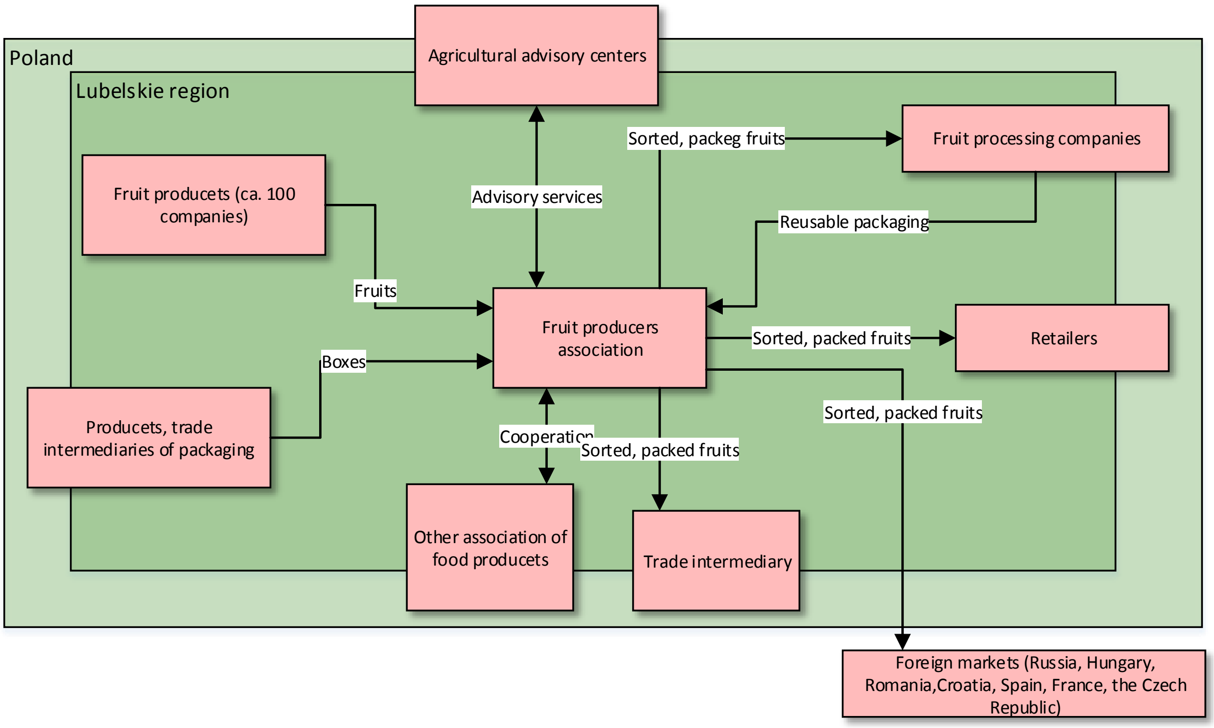


Figure 2 illustrates potential entities that could established business cluster initiative, which would be based on real cooperation networks and thus assure higher sustainability of such structure (in comparison to cluster identified purely with the use of location quotient).

3.2 Further work

Since the model was tested only on one industry in one region, further work will be focussed on three axes:

1. Substitution of the CATI data with Input-Output (I-O) analysis. Achievements in this area can lead to cost effectiveness, however its drawback is lack of availability of I-O in some countries on regional level. Even if such data is available in some of the EU countries, typically it is limited to the section level of NACE.
2. Development of visualization techniques of the business networks with the use of graphing software (with the use of Gephi or similar tool).
3. Practical implementation of the findings in cluster regional policies in Poland. The next funding period 2014-2020 poses real challenge to efficient support of the cluster initiatives, due to a number of unsuccessful projects in the preceding period. Failures are commonly associated with the lack of proper identification of cluster potential (focussing mainly on one sector, instead of value networks).

Acknowledgment

The authors are especially grateful to Marshal Office of the Lubelskie Voivodship in Lublin for providing data for our study. All remaining errors are the author's own.

References

Alonso-Villar, O., & Del Río, C. (2011). Concentration of Economic Activity: An Analytical Framework. *Regional Studies, 47*(5), 756-772. doi: 10.1080/00343404.2011.587796

Asheim, B., Cooke, P., & Martin, R. (2006). *Clusters and regional development : critical reflections and explorations*. London; New York: Routledge.

Asheim, B., Moodysson, J., & Tödtling, F. (2011). Constructing regional advantage: Towards state-of-the-art regional innovation system policies in Europe? *European Planning Studies, 19*(7), 1133-1139. doi: 10.1080/09654313.2011.573127

Boschma, R., & Iammarino, S. (2009). Related Variety, Trade Linkages, and Regional Growth in Italy. *Economic Geography, 85*(3), 289-311. doi: 10.1111/j.1944-8287.2009.01034.x

Brenner, T., & Muehlig, A. (2013). Factors and Mechanisms Causing the Emergence of Local Industrial Clusters: A Summary of 159 Cases. *Regional Studies, 47*(4), 480-507. doi: 10.1080/00343404.2012.701730

Bresnahan, T. F., & Gambardella, A. (2004). *Building high-tech clusters : Silicon Valley and beyond*. Cambridge, UK ; New York, NY: Cambridge University Press.

Crespo, J., Suire, R., & Vicente, J. (2014). Lock-in or lock-out? How structural properties of knowledge networks affect regional resilience. *Journal of Economic Geography, 14*(1), 199-219. doi: 10.1093/jeg/lbt006

Fornahl, D., Henn, S., & Menzel, M. P. (2010). *Emerging clusters: Theoretical, empirical and political perspectives on the initial stage of cluster evolution*: Edward Elgar Publishing.

Hidalgo, C. A., Barabási, A. L., Winger, B., & Hausmann, R. (2007). The product space conditions the development of nations. *Science, 317*(5837), 482-487. doi: 10.1126/science.1144581

Lechner, C., & Dowling, M. (2003). Firm networks: external relationships as sources for the growth and competitiveness of entrepreneurial firms. *Entrepreneurship and Regional Development, 15*(1), 1-26. doi: 10.1080/08985620210159220

Lechner, C., Dowling, M., & Welpe, I. (2006). Firm networks and firm development: The role of the relational mix. *Journal of Business Venturing, 21*(4), 514-540. doi: 10.1016/j.jbusvent.2005.02.004

Lehtonen, O., & Tykkyläinen, M. (2014). Estimating Regional Input Coefficients and Multipliers: Is the Choice of a Non-Survey Technique a Gamble? *Regional Studies, 48*(2), 382-399. doi: 10.1080/00343404.2012.657619

Rubach, S. (2013). Collaborative Regional Innovation Initiatives: A Booster for Local Company Innovation Processes? *Systemic Practice and Action Research, 26*(1), 3-21. doi: 10.1007/s11213-012-9270-8

Ter Wal, A. L. J., & Boschma, R. (2009). Applying social network analysis in economic geography: framing some key analytic issues. *Annals of Regional Science, 43*(3), 739-756. doi: 10.1007/s00168-008-0258-3

Walter, J., Lechner, C., & Kellermanns, F. W. (2007). Knowledge transfer between and within alliance partners: Private versus collective benefits of social capital. *Journal of Business Research, 60*(7), 698-710. doi: 10.1016/j.jbusres.2007.01.026

Contact

Maciej Piotrowski
University of Information Technology and Management
Sucharskiego 2 st., 35-225 Rzeszów, Poland
mpiotrowski@wsiz.rzeszow.pl

Korneliusz Pylak
Lublin University of Technology
Nadbystrzycka 38d St., 20-618 Lublin, Poland
korneliusz.pylak@pollub.pl

Michał Thlon
The Cracow University of Economics
Rakowicka 27 St., 31-510 Kraków, Poland
thlonm@uek.krakow.pl