

Chapter 6

Conceptual and Methodical Research Procedures



Miroslav Zizka , Petra Rydvalova , and
Vladimira Hovorkova Valentova 

6.1 Research Motivation

The published literature on economics has not paid enough attention to the analyses of the cluster existence impact on the performance of individual entities in a broader context—innovative, financial, and organisational. Based on the general equilibrium theory, it can be assumed that the existence of a cluster evokes both positive and negative microeconomic externalities in companies that also operate outside the industry. As far as the economic policy institutions are concerned, a description and understanding of the effects of the supporting instruments and programmes is essential for them. They enable to minimise negative and stimulate positive externalities of microeconomic policies. At the macroeconomic level, it is desirable to check whether cluster support in the form of subsidies is justifiable; see Pavelková and Jirčíková (2008).

The research carried out at the Technical University of Liberec within the Czech Science Foundation project from 2018 to 2020 sought to reveal the specific effects that clusters have on shifting the production function of all the companies involved as well as the entire industry. We can assume effects related to the improvement of the internal organisation of a company due to the existence of a cluster, and, subsequently in general, technological innovations realised as a result of more intensive cooperation of all entities in the industry. Production functions were modelled through linear programming, using data envelopment analysis (hereinafter DEA). Unlike other studies, the research focused not only on cluster organisations (hereinafter COs) established as a result of cluster initiatives, but also examined the effects of natural industry clusters in regions (so-called natural clusters). The

M. Zizka (✉) · P. Rydvalova · V. Hovorkova Valentova
Technical University of Liberec, Liberec, Czech Republic
e-mail: miroslav.zizka@tul.cz; petra.rydvalova@tul.cz; vladimira.valentova@tul.cz

research hypothesised that the existence of natural clusters could be a more effective tool in supporting corporate innovation activities than an artificially created CO.

As previously observed, even though institutionalised COs in the Czech Republic are supported from the European and national sources, the resulting clusters are not very successful in joining other Knowledge and Innovation Communities (KICs); see Pittnerová and Rydvalová (2014). This leads to the question of whether cluster initiatives really support important industries in the Czech Republic. Previous research on the functioning of clusters in the Czech Republic primarily focused on the innovative activities of COs. The team of Pavelková (2013) concluded that, based on a survey of 1110 member companies in the Czech clusters, the innovative activities of companies in the COs are at a low level. Higher values were reflected only in the manufacturing industry.

In evaluating the excellence of the Czech clusters, their role in the innovation environment was also examined in terms of cooperation with universities and research institutes, as well as in terms of companies' participation in innovation projects, economic data, and the results of research, development, and innovation. The results obtained led to the conclusion that clusters achieve better results after several years of development (the minimum of 3 years since their establishment). The evaluation of the innovation role of clusters was positively influenced by the higher share of universities and research centres in the membership base and their participation in innovative projects.

Rydvalová and Pittnerová (2013) analysed corporate innovation activities in the glass and bijouterie industry in the Czech Republic which have, especially in North Bohemia, the character of a natural cluster. The analysis showed that these companies mostly relied on marketing innovations. They mainly used private sources to finance these innovations. However, the overall companies' innovation potential was very low. It was revealed that the primary priority of these companies was to cope with their current problems, and their thoughts on future development, including innovation, were often postponed. Further research by Žižka and Rydvalová (2014) did not confirm the dependence of the intensity of the innovation in the regions on the number of clusters, their average age, or the size of their membership base.

The above-mentioned research was based on data covering a period of around 5 years. Czech COs have a relatively short tradition. Their development can be divided into two phases. In the first phase from 2002 to 2006, the concept of clusters was described in the Czech literature, and it was presented to the state administration and business managers in industries that were considered to have cluster potential. At the same time, mapping and clustering was supported through cluster initiatives with support from the European Structural Funds. Subsequently, over the period of 2004 to 2007, 53 cluster projects were supported from public funds, which mainly included support for mapping the possibilities of creating COs. In the second phase, which began in 2007 and has continued to the present day, support was focused on cluster activities aimed at innovation development and international competitiveness (Skokan et al., 2012).

This means that the oldest institutionalised cluster organisations existed for a maximum of 10 years at the time of the published outputs around 2013. Thus, it was

assumed that in this relatively short period of time published research results were influenced by too short a time period, during which it was impossible to demonstrate the positive impact of the CO existence. If the research is undertaken again after approximately five years (compared to the above-mentioned studies), the degree of validity of the results will increase.

6.2 Methodical Research Procedure

The research was divided into the following research activities; see Fig. 6.1.

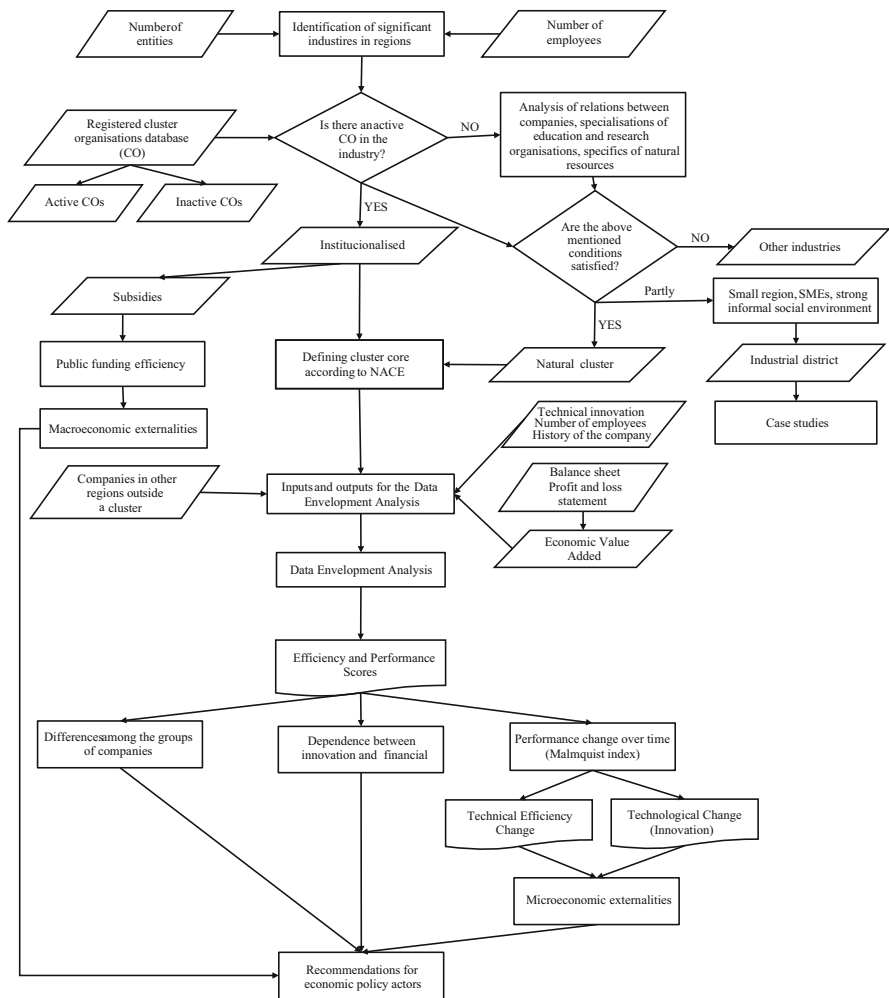


Fig. 6.1 The framework of the research

1. Identifying the Regions in the Czech Republic with Important COs The identification was performed based on the data from the Czech Statistical Office (hereinafter CZSO) register of economic entities on the number of these entities in individual districts of the Czech Republic. As of 31 May 2018, they were classified according to the CZ-NACE industry classification (three-digit codes). The second group comprised data on the number of employees by districts, obtained on 31 May 2018, classified again according to the CZ-NACE industry classification (two-digit codes). It is important to note that the data in both the sets is complete for the entire Czech Republic, which means that it is a statistical population, so it is not necessary to use statistical induction methods (hypothesis testing or statistical estimates) when processing the data.

When processing data on the number of economic entities in the districts of the Czech Republic, it can be assumed that if natural clusters are not formed in the given districts, in the given industry, the number of economic entities in them should be evenly distributed throughout the Czech Republic. However, it is necessary to take into consideration the fact that the size of individual districts varies, regardless of any comparison indicators (area, population, number of economically active people, etc.). As Brenner (2006) finds, the size of a region does not have a significant effect on the formation of clusters or the subsequent comparison of individual regions.

Therefore, the total number of economic entities in a given district was chosen as the distinguishing criterion weight. For each district, the proportion of the total number of economic entities and the total number of economic entities in the Czech Republic was calculated (Proportion 1). This expresses the assumption of an even distribution of the number of economic entities in the given industry across all the districts of the Czech Republic. Furthermore, the proportion of the number of economic entities in a specific industry in a given district and the number of economic entities in a given industry in total was determined (Proportion 2). Proportion 2 thus expresses how a given district contributes by the number of economic entities in a given industry to the total number of economic entities that are registered in this industry throughout the Czech Republic. Subsequently, Proportion 1 and Proportion 2 were compared in the way that Proportion 1 was subtracted from Proportion 2. If the result of this comparison is positive, it means that the actual proportion of the number of economic entities in the industry is higher in the given district than when an even distribution of the number of economic entities in the industry is presumed. Therefore, it can be implied that natural clusters could occur.

The same procedure was applied when processing the data on the number of employees in the districts of the Czech Republic. If there is no indication of the formation of a natural cluster in an industry, the number of employees in the industry should be evenly distributed across all the districts of the Czech Republic. In order to take into account different sizes of the districts in the Czech Republic, the proportion of the number of employees in the industry in the given district and the number of employees in the industry in the Czech Republic was established (Proportion A). Then the share of the current number of employees in the given industry in the

district and the total number of employees in the industry (Proportion B) was determined. Both Proportion A and Proportion B were compared in the way that Proportion A was subtracted from Proportion B. A positive result indicates the possibility of a natural cluster formation in the industry as the number of employees here is higher than in the distribution which would not presuppose the existence of a natural cluster.

Those industries which had a significant difference between Proportion 1 and Proportion 2 in a few districts were selected for further detailed analysis. This suggests that the activities in the given industry are concentrated in a limited space, and, therefore, it may signify activities of a natural cluster. The mathematical difference between Proportion B and Proportion A serves as a supporting evaluation criterion. Nevertheless, it does not provide as detailed information about the given industry as the difference between Proportion 2 and Proportion 1. The reason is that only data for the two-digit industry codes according to CZ-NACE were available.

2. Updating the Database of COs in the Czech Republic The input database was prepared as part of the team's previous research activities. COs were searched for using the public register and the administrative register of economic entities. All organisations that contained the keyword 'cluster' in their name were identified. The database comprised the name of the CO, its identification number, industry, legal form, year of establishment, list of members, number of employees, the region of operation, registered office site, contact details, and a link to the website.

The database contained data on 114 institutionalised COs in the Czech Republic, of which 16 already ceased their operation as of 1 January 2019. The remaining 98 organisations were further analysed in terms of their activities. There were identified 74 active COs with projects and up-to-date information on the cluster official website. Their financial statements in the public register and collection of documents can be traced. The remaining COs are inactive. They either do not have an official website or the website is down; thus, they do not have financial statements available in a public register and a collection of documents.

3. Identifying Other Potential Clusters in the Regions of the Czech Republic The identification is based on

- National Report on Clusters from 2006 (Adámek et al., 2006).
- Internal research based on the calculation of industry location quotient (Žižka, 2006).
- Concentration of the industry in the regions (see phase 1).
- Analysis of the relations among the companies.
- Presence of research and educational institutions (see Sect. 6.2.1).

The location quotients calculate the degree of employment concentration in the given industry and in the specified region. The data on employment by region were adapted from the CZSO regional statistics and the MagnusWeb commercial database (Bisnode). Financial data and information on links among companies were also obtained on the basis of the licence purchased. The output of this activity was to

determine the status of a given industry in the region (a natural cluster, a CO, an industrial district, other—minor industries), i.e. a typology of the industry.

Based on the intertwined results of phases 1–3, the regions were identified in which both natural clusters and institutionalised COs exist. This situation occurred, for example, in the NUTS 2 Northeast region, where there is a natural group of textile manufacturers and a CLUTEX CO (focusing on technical textiles). At the level of NUTS 3, these are, for example, clusters (natural and institutionalised) of packaging material producers in the Hradec Kralove Region, food and timber grouping in the Jihočeský Region (the South Bohemian Region), timber and energy clusters in the Moravskoslezský Region (the Moravian-Silesian Region), an electrical engineering cluster in the Pardubice Region, an IT cluster in the Vysočina Region, as well as a packaging cluster in the Zlín Region.

Some institutionalised clusters connect entities operating in several regions of the NUTS 3 level. *It is complicated to define natural clusters which emerged historically and whose borders do not copy the administrative division of the regions in the Czech Republic.* For this reason, based on a research case study method, authors' own procedure that was tested in the conditions of the Czech Republic was prepared (see Sect. 6.2.1).

For further research, seven randomly selected industries were chosen, which had to fulfil the condition that there was a CO or a natural cluster (see Table 6.1). In the case of a CO, the condition had to be satisfied that it was already in its maturity stage (established by 2010, at the latest). Furthermore, two industries in which there is only a natural cluster (glass and bijouterie industry) were added to the research. The hypothesis was also verified that the COs were only established in regions with a significant representation of the given industry (see Sect. 6.2.2).

4. Identifying the Core of Clusters For both natural clusters and COs, it was necessary to define the companies that make up the cluster core. These are companies from a given industry or related industries that can be considered homogeneous in terms of production inputs and outputs. *The research focused on business entities, not on public or non-profit institutions which are usually members of clusters as well.* The MagnusWeb database, a collection of documents and a public register that contains information on the main business activity of each entity according to the NACE statistical classification, served as a source of data. A list of rated companies for each group was compiled; *each company could only be on one list.* If a company is in an institutionalised cluster (CO), which also logically operates in the territory of the natural cluster, then it was included only in the list of the CO. The procedure for determining the economic entities of a natural cluster core is described in Sect. 6.2.1.

5. Preparing a List of Other Companies A list of companies operating in other regions (outside institutionalised or natural clusters) was also prepared for all the industries covered by the research. The list served as a control group for comparing performance with the clustered companies.

6. Collecting Data on Technical Innovations (According to the Oslo Manual) Data on protected results under industrial law, such as patents, utility models,

Table 6.1 Characteristics of the analysed industries

Industry	CO's region of operation	Core NACE	Established in	No. companies in CO	No. companies in the CO's region except of CO members	No. companies outside the CO's region
Automotive	Moravian-Silesian	293	2006	11/9	25/13	135/52
Engineering	Moravian-Silesian, Olomouc, South Bohemian	251, 28	2003	10/6	1011/167	1836/197
IT	Moravian-Silesian	620	2006	8/6	718/21	3409/85
Furniture	South Bohemian	161, 162, 310	2006	16/12	1794/46	9125/68
Nanotechnology	Central Bohemian, Pardubice	721	2010	6/3	357/19	790/45
Packaging	Hradec Kralove	172, 222	2005	16/9	104/15	764/162
Textile	Liberec, Hradec Kralove, Pardubice	13, 141	2006	19/19	275/38	1430/138
<i>Total</i>				<i>86/64</i>	<i>4284/319</i>	<i>17,489/747</i>

Source: own processing (2020)

industrial designs, and trademarks, were collected for all the companies in the examined industries. For this purpose, an extensive search in the database of the Industrial Property Office of the Czech Republic was performed, and the number of the above-mentioned industrial property rights was ascertained for all the companies listed in phases 4 and 5. Furthermore, information on the licensing of industrial property rights was also added when available.

7. Collecting the Data from the Balance Sheets and Profit and Loss Statements The accounting data necessary for the calculation of economic value added and a subsequent DEA analysis were obtained for all companies identified in phases 4 and 5. The MagnusWeb database and a collection of documents from the Commercial Register serve as the source of data. The data obtained cover the period of 2009–2016.

The basic characteristics of the selected COs are listed in Table 6.1. The number of companies in the relevant groups is displayed first, followed by the numbers of companies for which financial data were obtained in the entire time series 2009–2016. The number of companies analysed is influenced by the fact that sole trader businessmen are not obliged to publish financial statements. Business corporations in the micro and small enterprises category publish their condensed financial statements without a profit and loss statement, which also impedes in-depth analyses of their financial performance.

These two categories of companies are most highly represented in the economy. According to the CZSO (2019), 97.5% of entities in the manufacturing industry had less than 20 employees in 2018. However, even some larger corporations do not comply with legal obligations and do not publish their accounting data, or they do so with a significant time lag. As Table 6.1 shows, it was easiest to get accounting data on the CO member companies. Most of them were business corporations that have to publish annual reports and financial statements in the public register.

8. Collecting the Data on the Number of Employees and the History of the Companies Accounting data were supplemented by the number of employees, which is a significant input for further evaluation of technical efficiency and performance of the companies. The data were obtained from the MagnusWeb database and public registers. The history of a company can be considered a form of accumulated intellectual capital containing a technological trajectory, routines, know-how, skills, and experience of the owners and their employees. The history was identified based on the data in the public register of the Czech Republic and information on the companies' websites. The information about family businesses was further added for an easier identification. This factor (see RQ10 in the introduction) is assumed to mainly appear in natural clusters.

9. Obtaining Information on Subsidies from Public Funds Spent on the Creation and Development of Cluster Initiatives The CEDR information system of the Ministry of Finance of the Czech Republic (hereinafter MF) was used as a data source. The information was obtained for the above-mentioned seven COs.

10. Economic Value-Added (EVA) Calculations These were performed for companies with available financial statements. EVA is based on economic profit, which counts all the costs of the invested capital, both company's own and borrowed (Grant, 2003). EVA is considered a modern measure of a company's success because it expresses the true profitability of the company and is associated with the requirement to maximise shareholder wealth (Stewart, 1994). As the transformation of accounting profit into economic profit is a relatively complicated process, also influenced by national accounting standards, Neumaierová and Neumaier (2002) created a model for calculating the EVA indicator; see Eq. (6.1). The same calculation methodology has been used by the Ministry of Industry and Trade since 1999. This is an equity-based process in which EVA is defined as the product of equity and spread (i.e. return on equity minus the alternative cost of equity):

$$EVA = (ROE - r_e) \cdot E \quad (6.1)$$

ROE Return on equity
r_e alternative cost of equity
E Equity

The alternative cost of equity (*r_e*) can be calculated using formula (6.2), where a risk premium is added to the risk-free rate (*r_f*). According to the MPO (2017), the risk premium consists of a risk premium for business risk (*r_{bus}*), financial structure (*r_{finstr}*), financial stability (*r_{finstab}*), and the size of the company or the liquidity of its shares (*r_{ls}*).

$$r_e = r_f + r_{bus} + r_{finstr} + r_{finstab} + r_{ls} \quad (6.2)$$

The EVA value was determined both individually for each business entity and for entire groups of companies (a CO, non-member companies operating in the same region as the CO, companies outside the COs region) based on the aggregation of company data.

11. Evaluating the Innovation and Financial Performance of Companies Using Data Envelopment Analysis (DEA) This phase included the definition of inputs and outputs, which should be independent (Düzakın & Düzakın, 2007). The relation between an input and an output (within and among the groups) was examined using correlation analysis. Assets, long-term capital, number of employees, duration of the existence of the company, etc., can be considered inputs for the evaluation of innovation activities. Outputs include patents, utility models, industrial designs, and trademarks; also the commercial success of patenting in the form of a licence is advantageous.

The output of the first phase can serve as an input for the second phase of the evaluation that examines how businesses can commercially use protected results.

Economic indicators (EVA, sales, and revenue to cost ratio) served as outputs from the second phase. After determining the suitable inputs and outputs, the formulation of mathematical models (under the conditions of constant returns to scale; hereinafter CRS and variable returns to scale; hereinafter VRS) was created and the results were calculated by using the OSDEA-GUI and MaxDEA Ultra software tools. The efficiency score in the innovation and financial areas and the overall performance score were defined for each company.

The aim of the DEA model is to maximise the objective function z (6.3) under constraints (6.4). The inputs x_j of the unit q have weights v_j . The outputs y_i have weights u_i . Units that are efficient have an objective function value equal to one. These units are located at the efficient frontier. Inefficient units have value z lower than one.

$$z = \sum_{i=1}^r u_i y_{iq} \tag{6.3}$$

$$\begin{aligned} \sum_{i=1}^r u_i y_{ik} &\leq \sum_{j=1}^m v_j x_{jk}, \quad k = 1, 2, \dots, n \\ \sum_{j=1}^m v_j x_{jq} &= 1 \\ u_i &\geq \varepsilon, i = 1, 2, \dots, r; \\ v_j &\geq \varepsilon, j = 1, 2, \dots, m \end{aligned} \tag{6.4}$$

In the case of variable returns to scale, it is sufficient to add to the model the variable μ indicating the deviation from the CRS. In this case, the conical data envelope changes to convex, which leads to a higher number of efficient units being defined. The input-oriented BCC model has the form given by relations (6.5).

$$\begin{aligned} z &= \sum_{i=1}^r u_i y_{iq} + \mu \\ \sum_{i=1}^r u_i y_{ik} + \mu &\leq \sum_{j=1}^m v_j x_{jk}, \quad k = 1, 2, \dots, n \\ \sum_{j=1}^m v_j x_{jq} &= 1 \\ u_i &\geq \varepsilon, i = 1, 2, \dots, r; \\ v_j &\geq \varepsilon, j = 1, 2, \dots, m; \\ \mu &\in R \end{aligned} \tag{6.5}$$

12. Identifying the Differences between the Individual Groups of Companies within Each Industry This was performed by using the Kruskal–Wallis test. The test is a non-parametric variant of the variance analysis that tests differences between median performance scores. Furthermore, the differences in the means of the individual groups were examined by using the Games-Howell posthoc test. It is a non-parametric approach to comparing multiple groups of observations which works with the order of the original values. Its advantage is that it does not require compliance with the conditions of a normal distribution, homogeneity of variance, or the same group size. All tests were performed at the significance level of alpha 5%.

13. Examining the Relation Between the Performance and Innovation Activities of Companies The performance was measured by using a performance score which was the result of the DEA models solved separately within individual industries and subgroups of the companies. The value of the performance score was converted into a binary variable with values of 0 or 1. The value of 1 was assigned to companies that were marked as performance units or best practices in the group. Other companies, with the performance score lower than 1, were given a value of 0. Using the Chi-square test of independence for categorical variables, the connection between the registered results of innovation activities and the status of a high-performance unit was examined. The strength of the dependence was measured using Pearson's correlation coefficient R .

14. Evaluating Performance Changes During the time line of 2009–2016, this evaluation was performed using the Malmquist index. The Malmquist index (hereinafter MI) evaluates the changes in relative productivity or performance of a decision-making unit between different time periods. One of the advantages of MI is the identification of the components that lead to a performance change. The index breaks down total productivity change into technical efficiency change EFFCH and technological change TECH; see Eq. (6.6). Companies aim to approach the closest to the efficiency frontier (the best companies in the industry) with the help of various internal organisational measures. The EFFCH component expresses this effort (Li et al., 2017). Simultaneously, however, the innovations in the industry cause the efficiency frontier shift over time. The TECH component expresses this shift. In general, it is desirable for the MI, EFFCH, and TECH values to be greater than one. In that case, industry productivity increases, efficiency improves, and technological progress takes place.

The internal technical efficiency change (EFFCH) can be further decomposed into the product of pure technical efficiency change (PECH) and scale efficiency change (SECH). PECH expresses the company's ability to improve its internal technical efficiency between two time periods, t and z , under the conditions of VRS. SECH measures the change in scale efficiency between these periods. The optimal value of SECH is unitary as, in this case, the company operates in conditions of the CRS and produces in the technically best range (Pantzios et al., 2011).

$$MI^{t,z}(x^z, y^z, x^t, y^t) = EFFCH_1TECH_1 = (PECH_1SECH_1)TECH_1 \quad (6.6)$$

Suppose a company that works in period t with a vector of n inputs x^t and produces m outputs y^t . Then (x^t, y^t) represents the input–output pair of the given company in the period t , and (x^z, y^z) is the input–output pair of the same unit over the time period z . The Malmquist index between the time periods z and t is represented by the equations in (6.7) and (6.8).

$$MI^{t,z}(x^z, y^z, x^t, y^t) = \sqrt{\frac{D^t(x^z, y^z)D^z(x^z, y^z)}{D^t(x^t, y^t)D^z(x^t, y^t)}} \quad (6.7)$$

$$\begin{aligned} MI^{t,z}(x^z, y^z, x^t, y^t) &= EFFCH^{t,z} \times TECH^{t,z} \\ &= \frac{D^z(x^z, y^z)}{D^t(x^t, y^t)} \times \sqrt{\frac{D^t(x^z, y^z)D^t(x^t, y^t)}{D^z(x^z, y^z)D^z(x^t, y^t)}} \end{aligned} \quad (6.8)$$

The expressions $D^t(x^t, x^t)$ and $D^z(x^t, x^t)$ are called distance functions. They express the distance between the input and output units over the time period t and efficient frontiers over the time periods t and z . Another pair of distance functions $D^{t,z}(x^z, x^z)$ measures the distance between input and output units over the time period z and efficient frontiers over the time periods z and t (Wang, 2019). The values of distance functions D are estimated with the help of DEA.

Equation (6.8) shows the decomposition of MI into two components. The change in the technical efficiency $EFFCH^{t,z}$ assumes an efficient frontier in the same period and expresses the ratio of the efficiency of a given company over the period z to its efficiency over the period t . It shows how a unit tried to improve its internal performance through various measures in work organisation or production. The technological change $TECH^{t,z}$ characterises the frontier shift. If the efficient frontier shifts from a position over the time period t to a position over the time period z , the value of $TECH^{t,z}$ will be greater than one and technological progress will occur. The component $TECH^{t,z}$ expresses the group change in the efficiency caused by all the companies, i.e. innovations in the industry (Wang, 2019).

15. Comparing the Malmquist Index and its Components among the Groups of Companies and Industries As the final output of the previous phase was geometric averages, the original company values of MI and its components were logarithmised. The geometric mean is a monotonic function of the mean of logarithms. If there is a significant difference between the means of the logarithmised data, there is also a significant difference between the geometric means of the original variables (Alf & Grossberg, 1979).

The Shapiro–Wilk test showed that the data did not have a normal distribution, and in some groups of companies the Levene’s variance check test found that the

condition of homo-scedasticity was not satisfied. For these reasons, a non-parametric Games-Howell posthoc test was used to identify all the differences among all the company groups. The Games-Howell test works with the order of the original values and examines the differences in means of individual groups. Statgraphics XVIII software was used to test the differences. The alpha significance level was 5%.

A statistical hierarchical cluster analysis was performed to identify the similarities among individual groups of companies in different industries. Ward's method of clustering with squared Euclidean distances was used. The similarity of groups of companies was first assessed using a dendrogram of objects. Furthermore, the basic characteristics of MI and its components EFFCH and TECH were calculated for each statistical cluster.

16. Identifying and Formulating Macroeconomic and Microeconomic Externalities of a Company Clustering The question of whether the existence of clusters, especially the institutionalised COs, is a positive macroeconomic externality and whether it should be supported by economic policy actors was addressed. To answer this question, a literature review was carried out. The empirical verification of the positive externalities was done by analysing the additional tax and non-tax revenues in public budgets of the seven selected COs. The increase in corporate income tax of the clustered member entities was evaluated, as well as the increase in personal income tax of their employees, along with the increase in social and health insurance, paid by both companies and employees.

The benefits of sharing the results in research, development, innovation, and tacit knowledge, as well as the multiplication of knowledge leading to increased economies of scale, are considered the positive microeconomic externalities of clustering. The data obtained from the literature review were empirically verified by the research on the companies' approach to innovation in each industry, and changes in internal technical efficiency of the companies.

17. Synthesising the Outputs, Formulating the Consequences, and Recommendations for the Corporate Inter-organisational Behaviour Based on the data obtained, recommendations were formulated for the actors of the economic policy on how to approach support of clusters. The results identified were also processed via case studies for the textile, glass, and bijouterie industries. The case studies mentioned demonstrate the influence of other selected factors, such as tradition, or the type of innovation in the family businesses, on the development of the industries in which they create natural industrial districts and clusters in the regions studied.

6.2.1 Procedure for Defining Natural Cluster Cores

Due to further research, it is necessary to distinguish two basic types of industry clusters.

The first type is those clusters that arose naturally in compliance with Porter's definition, without any external state intervention. As mentioned, these are referred to as *natural clusters*. A natural cluster is a cluster of interconnected companies and research and educational institutions, which has existed in the given area for a long time, and no public resources were spent on its initiation.

Then, there are COs that are the result of a certain cluster initiative. A cluster initiative is an organised effort to increase the growth and competitiveness of clusters in the region, involving the cluster companies, the government, and the research community (Lindqvist et al., 2012). The umbrella CO then provides the grouping with certain management services across the business and innovation processes. A CO can exist as a subset of a natural cluster as it generally does not include all industry organisations that operate in a region where a natural cluster exists. It is important to define the criteria for determining the existence of the natural cluster as this can help in evaluating the performance of both types of clusters.

It can be assumed that natural clusters, which are not formalised and do not have contractual partnerships, will solve different problems than organised clusters, which are often set up to obtain some public support for their further development. In the case of industry clusters which arose based on the historically existing resources (labour, land, capital) in a given region, it can be stated that they have strong roots in the given region.

The question is how to identify such a region, how to find out the importance of a given industry for its development, and, in particular, how to identify specific economic entities of a given natural cluster. Natural clusters must necessarily copy the administrative boundaries of the region to which the available data from the statistical office are linked. It is often necessary to first find the boundaries of a functional region which are defined by the local labour market and commuting areas; see Žižka (2013). Porter (1998) states that geographic borders of a cluster can vary from cities to states, to a group of neighbouring countries. Therefore, the problem is not to identify key, upstream, and downstream industries but to define geographic borders of a natural cluster. COs, established with the help of public resources, were geographically defined according to the administrative region borders in compliance with the statistical classification NUTS (NUTS 2 or NUTS 3).

To define a natural cluster, pilot research was first conducted, using the method of a research case study based on local knowledge. The methodical procedure of setting up a research case study falls into the field of qualitative research including selecting a data collection method, choosing a research sample, creating a case study protocol for the database, contacting the research subjects, processing the data in the database, analysing and evaluating the obtained data, comparing the theoretical knowledge with the findings, and formulating research questions (Štrach, 2007). Based on the previous study (Rydvalová & Hotař, 2012), the glass and bijouterie industry was chosen for the pilot testing. The given industry was selected for several reasons. It is an industry which is historically typical for North Bohemia, and the authors have experience in cooperation with economic entities in this industry (manufacturing, educational, consulting, research). Furthermore, in the previous research, they

developed a map of links among the entities in this industry (upstream, core, follow-up activities, including the infrastructure).

It is common in studies mapping the existence of clusters (e.g. Delgado et al., 2014) to define industries with four-digit codes SIC (Standard Industrial Classification—an American system classifying industries). Such accuracy is not possible in the Czech Republic because it encounters the issue of the General Data Protection Regulation. Therefore, the following procedure for the identification of a natural cluster was used in a case study of the glass and bijouterie industry in North Bohemia.

The first step was to identify the dominant industry within the geographically defined region. Ideally, municipalities would serve as the basis for industry analysis of data at the national level for the whole republic. Nevertheless, the Czech Statistical Office regards the data on the number of entities and the number of employees at the municipal level as individual data. The data on the number of economic entities (according to the NACE classification at the level of three-digit codes) and the number of employees (according to the NACE classification only at the level of two-digit codes) are thus available in the aggregate form only at the district level. Districts (NUTS4/LAU1) in the Czech Republic are relatively heterogeneous in terms of size (area, population, number of economic entities). The procedure for identifying significant industries in the districts was as follows.

The proportion of the number of economic entities in individual districts in the total number of economic entities in the Czech Republic was determined. This proportion characterises the economic size of the district. For example, in Prague, there are 21.42% of all economic entities; the smallest number of them are in the Jeseník district, totalling only 0.37%.

- (a) The actual proportion of economic entities in a given district and industry (to the three-digit NACE codes) was compared with the proportion calculated in the previous step. If the actual proportion of the number of economic entities in the industry, in the given district, and the number of all economic entities in the given industry is greater than the one that expresses, the size of the district is, de facto, defined by the number of economic entities, thus indicating the possibility of the existence of a natural cluster. The difference need not be tested statistically; a basic set of all economic entities is sufficient for the needs of the analysis. The results for NACE 231 (manufacture of glass and glass products) and NACE 321 (manufacture of jewellery, bijouterie, and related articles) for selected districts are displayed in Table 6.2.
- (b) By analogy, the proportion of employment in the industry, in the given district, and the proportion of the district in the total employment were calculated, and the difference between the two variables was established; see Table 6.3. The fact that the data on employment are only available at the level of the two-digit NACE codes can be seen as a disadvantage. Thus, in many cases, they may cover a wider range of economic activities than in the case of the previous analysis by the number of economic entities. For example, the industry NACE 23 (manufacture of other non-metallic mineral products) includes, in addition to

Table 6.2 The concentration of economic entities in the CZ-NACE 231 and 321 industries, in selected districts in the Czech Republic

Industry CZ-NACE 231: Glass production and glass articles			
District	The proportion of the district on the number of entities in the industry	The proportion of the district on the number of economic entities in the Czech Republic	Difference
CZ0511 Ceska Lipa	0.1988	0.0081	0.1907
CZ0512 Jablonec n. N.	0.1720	0.0085	0.1636
CZ0631 Havlickuv Brod	0.0421	0.0071	0.0350
...			
Industry CZ-NACE 321: Manufacture of jewellery, bijouterie, and related articles			
CZ0512 Jablonec n. N.	0.1491	0.0085	0.1406
CZ0514 Semily	0.0530	0.0069	0.0460
CZ0513 Liberec	0.0369	0.0178	0.0191
...			

Source: own processing according to CZSO (2018)

glass and glass product manufacture, the manufacture of refractory, building, porcelain and ceramic products, cement, lime, plaster, and abrasive products, etc. NACE 32 represents other manufacturing, apart from the manufacture of jewellery and bijouterie, as well as manufacture of musical instruments, sports goods, games and toys, medical and dental instruments and supplies, etc. For this reason, the analysis of the number of entities was chosen as a determining factor for identifying natural clusters, and the analysis of the number of employees was considered complementary. The effect of the accuracy of the industry category is evident in the case of small industries. Table 6.2 indicates that entities in the bijouterie industry are mostly concentrated in the Jablonec nad Nisou district. However, this fact is difficult to notice in Table 6.3.

- (c) As regards employment, other manufacturing is mainly concentrated in the Kladno district. A further examination reveals that out of almost 3800 employees in this broad industry, approximately 2500 are employed by Lego Kladno, the manufacturer of games and toys.

In the second step, only those industries that are significantly concentrated (especially in terms of the number of entities) in individual districts were analysed. First, it was determined whether there was a CO established based on a cluster initiative in the given region and industry. The list of COs was created as a part of our research (see Sect. 6.2).

Table 6.3 The concentration of employees in the CZ-NACE 23 and 32 industries in selected districts, in the Czech Republic

Industry CZ-NACE 23: Manufacture of other non-metallic mineral products			
District	Proportion of the district on the employment in the industry	Proportion of the district on the number of employees in the Czech Republic	Difference
CZ0423 Teplice	0.6389	0.0232	0.6157
CZ0512 Jablonec n. N.	0.3900	0.0171	0.3729
CZ0533 Svitavy	0.2473	0.0186	0.2287
...			
Industry CZ-NACE 32: Other manufacturing			
CZ0203 Kladno	0.1022	0.0096	0.0926
CZ0311 Ceske Budejovice	0.0746	0.0195	0.0551
CZ0521 Hradec Kralove	0.0679	0.0147	0.0532
...			

Source: own processing according to CZSO (2018)

To identify natural clusters, we are interested in those industries where there are no institutionalised COs; they will be evaluated separately. In the case a given industry shows a significant concentration in the region (in terms of the number of entities and employees), and, at the same time, there is no support for a CO existence, then it is a candidate for creating a natural cluster. The subsequent analysis evaluates the number and profiling of the entities in the given industry in municipalities of the given region.

In the third step, the fulfilment of other factors necessary for the existence of a natural cluster in the given region (the district level was chosen/NUTS4/LAU1) was checked by monitoring the publicly available data on:

- The specialisation of the vocational and higher education (source: Ministry of Education Youth and Sports, hereinafter MYES—register of schools).
- Industrial traditions in the region, including the existence of traditional and family businesses, monitoring customs, tacit knowledge, and skills. This factor is very demanding on the knowledge of the local environment, as there is no complete material mapping the historical development of individual industries in the Czech Republic, yet a study conducted at Masaryk University may help with this (Svobodová et al., 2013).
- The specialisation of the research organisations (data source: MYES—register of research organisations).

- Availability of local natural resources concerning the industry.
- Existing networking.

It should be noted that the analysis of industrial tradition and the structure of the region requires excellent knowledge of the local environment. Therefore, it is necessary to obtain data from local organisations. The data on natural clusters are further processed by research case studies. An example of such a situation is the ‘Crystal Valley’ natural cluster (Rydvalova & Zizka, 2021, in this book).

In the fourth step, the municipalities were specified, through which it was possible to define the region of the natural cluster. The method for evaluating a systematic and predictable lack of balance according to Vilfred Pareto’s rule was chosen to select the important municipalities for the analysis of economic entities with the identified industry according to the NACE in the district. As Pareto proved (Koch & Novotná, 2008), the measure of such inequality is the ratio 80/20, which points out to the fact that approx. 80% of the outputs are the results of approx. 20% of the inputs. The procedure can be summarised as follows:

- Preparation of data at the level of municipalities (the source of the data is the public register of the CZSO), the selection of the key industries at the two-digit NACE.
- Arranging the data in descending order according to the size of all the municipalities in the observed regions.
- Building a histogram (cumulative sum).
- Expressing the partial sums in the percentage of the total number.
- Setting the criteria for decision-making.
- Separation of the essential from the insignificant.

Graphically, it is represented by the Lorenz curve which expresses the degree of concentration of the observed phenomenon. It is a graph of quantiles related to the uniform distribution.

The evaluation focused on industries, which according to the classification of economic activities (NACE) were classified by Eurostat as key industries in terms of innovation activities, whereas all wholesale activities were excluded. The included industries are expressed by two digits:

- Mining and quarrying—B/05-09/.
- Manufacture of food products—C/10-34/.
- Electricity, gas, steam, and air conditioning supply—D/35/.
- Water supply, sewerage, waste management, and remediation activities—E/36-39/.
- Transformation and storage—H/49-53/.
- Information and communication—J/58-63/.
- Financial and insurance activities—K/64-66/.
- Architectural and engineering activities; technical testing and analysis—M/71/.
- Scientific research and development—M/72/.
- Advertising and market research—M/73/.

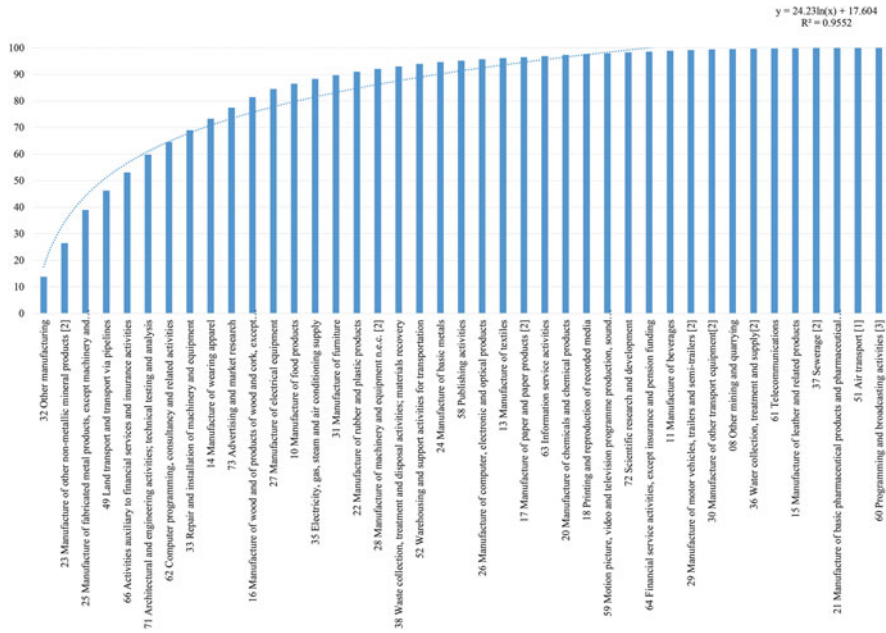


Fig. 6.2 The significance of knowledge industries in the Jablonec nad Nisou district, in compliance with CZ-NACE (two digits), excluding wholesale in %. *Source:* own processing according to CZSO (2018)

These industries were analysed in a pilot study in 34 municipalities of the Jablonec nad Nisou district, according to the economic entities indicator. As displayed in Fig. 6.2, NACE 32 and NACE 23 industries ranked in the first and second places according to the number of economic entities, followed by the industries which subsequently provide only services to the entities in these industries. Following these findings, the number of entities indicator was observed for individual municipalities only in the above-mentioned industries NACE 23 and NACE 32, and again, in compliance with Pareto’s rule, Therefore, the municipalities from the Jablonec nad Nisou district were chosen for an in-depth analysis of economic entities. These municipalities can be viewed as the core of the evaluated natural cluster from the geographical point of view.

The industry NACE 23: Zelezný Brod, Jablonec nad Nisou, Pencin, Smrzovka, Tanvald, Mala Skala, Koberovy, and Desna. The eight municipalities make up 24% of the municipalities bringing 80% of the economic entities in the given industry CZ-NACE 23 in the district (see Fig. 6.3).

Industry NACE 32: Jablonec nad Nisou, Zelezný Brod, Pencin, Smrzovka, Tanvald, and Rychnov u Jablonec nad Nisou. The six municipalities make up 15% of the municipalities bringing 80% of the economic entities in the given industry CZ-NACE 32 in the district (see Fig. 6.4).

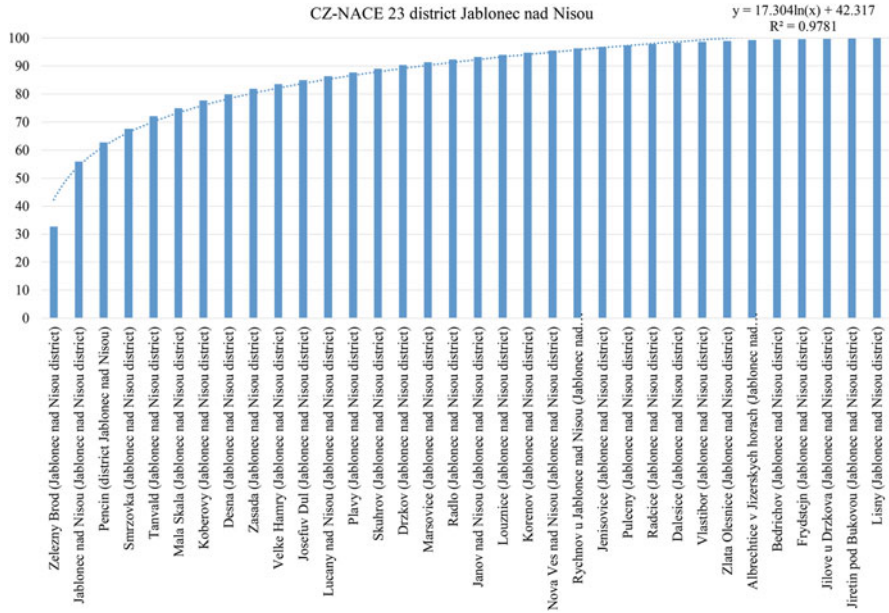


Fig. 6.3 The significance of municipalities according to the number of economic entities with CZ-NACE 23 in % in the Jablonec nad Nisou district

However, there is a certain risk that a significant entity in the municipality can be overlooked and not included in the analysis. The significance threshold must be intuitively set based on the knowledge of the local environment.

Therefore, for example, the municipality of Zasada must be included in the analysis of NACE 23 (after all, the number of economic entities (16) in Zasada does not differ so much from the number of entities in the municipality of Desna (18), included in the analysis automatically according to the above algorithm). In the case of NACE 32, the municipalities of Lucany and Mala Skala were added to the analysis for similar reasons.

In the fifth step, a detailed analysis of economic entities was conducted, and the data were obtained from the publicly available registers in the Czech Republic (ARES, 2020) in the above-identified municipalities and industries. Based on this, a database of economic entities was created. It enables a consequent evaluation within a natural cluster.

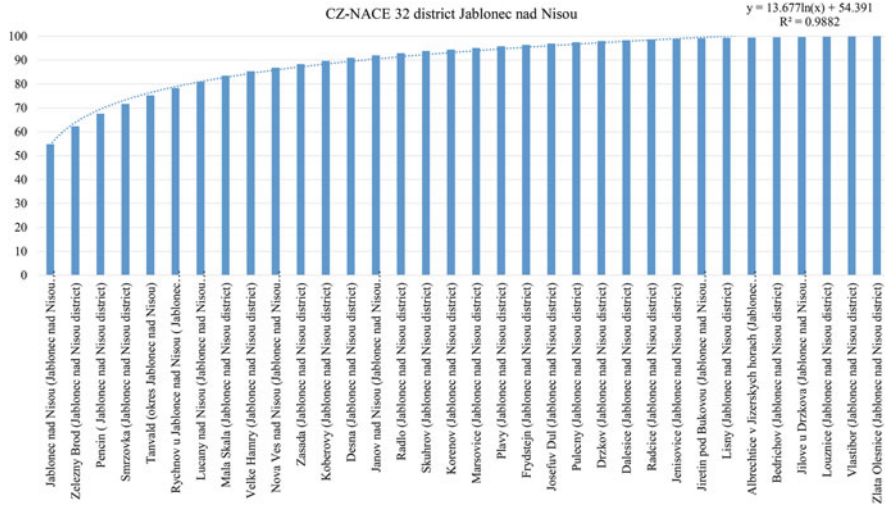


Fig. 6.4 The significance of municipalities according to the number of economic entities with NACE 32 in % in the Jablonec nad Nisou district. *Source:* for both figures: own processing according to CZSO (2018)

6.2.2 Methodological Procedure for Assessing the Territorial Distribution of Cluster Organisations

When assessing the territorial distribution of COs, the intention was to prove that COs prevail in districts or regions where there is a significant representation of the industry in which the CO operates. This means that the proportion of the members of a CO sample operating in a district/region, where the representation of the given industry is significant, must be greater than 50%. The process of preparatory and implementation work was scheduled as follows:

- Assigning a district and a region to the place of operation of each member of all the seven monitored COs.
- Assessing whether the place of operation of the CO members belongs to the district/region, in which there is a significant representation of the given industry.
- Calculation of the proportion of the CO members who operate in the district/region, where the industry is significant.
- Assessing individual COs in terms of their place of operation in the district/region with a significant representation of the given industry. The assessment was run by a simple comparison of the calculated proportions in the region/district with a value of 50%. The research was carried out for seven randomly selected clusters. A simple comparison was used because the authors de facto worked with seven populations, where the observed feature was checked in all their units.
- One sample was created out of all the members of the seven monitored COs. Two hypothesis tests for π were run, which are described in detail for example in Black

(2010). The reason for using this test was the fact that from the set of COs that were considered suitable for the analysis, these seven were randomly selected. Based on that, it is possible to work with this set as with one-stage cluster sampling, where during the first step a number for groups was chosen, and in the second step, all the units of the selected groups were assessed, as described in Lohr (2010).

Based on the above procedure, it was examined whether business entities that are members of cluster organisations have their registered office mainly in districts or regions, where a significant concentration of the industry is identified. Spatial concentration was measured using a localisation coefficient. The coefficient compares the relative employment in a given industry at the regional level to the relative employment in the industry at the national level. An industry is considered significant in a region if the value of the localisation coefficient exceeds 1.1 (Skokan, 2004).

At the district level, out of seven randomly selected cluster organisations, only two cluster organisations (CLUTEX and IT Cluster) were mostly located in areas with a significant concentration of the industry. In the case of regions, four such organisations were identified. On average, 51% of the members of institutionalised clusters were based in a region with a significant representation of the industry. These were IT Cluster, Czech Machinery Cluster, and MS Automotive Cluster and the CLUTEX cluster organisations. Subsequently, at the level of significance of 5%, it was tested whether the members of cluster organisations significantly occur in the territory where the given industry predominates, namely at the regional level. This hypothesis could not be proved.

It is obvious that, in fact, natural or functional regions may differ from administratively defined units. Therefore, it was determined whether the members of cluster organisations are located near the border of the administrative region, specifically within a distance of 40 km from it. This distance was selected because commuting radius is supposed to be within one hour. It was found that the proportion of members of cluster organisations within commuting distance increased to 74%. Subsequently, at a significance level of 5%, it was demonstrated that the proportion of members with the registered office within 40 km from a district where the industry is significant is higher than 50%. It can be concluded that Czech institutionalised clusters were indeed created in areas where the industry traditionally occurs.

References

- Adámek, P., Vlasák, T., & Stejskal, J. (2006). *Statistická identifikace klastrů: Národní zpráva—Shrnutí poznatků*. Berman Group.
- Alf, E. F., & Grossberg, J. M. (1979). The geometric mean: Confidence limits and significance tests. *Perception & Psychophysics*, 26(5), 419–421.
- ARES. (2020). *Administrativní registr ekonomických subjektů – ekonomické subjekty*. Ministry of Finance. https://www.info.mfcr.cz/ares/ares_es.html.cz

- Black, K. (2010). *Business statistics: For contemporary decision making* (6th ed). Wiley.
- Brenner, T. (2006). Identification of local industrial clusters in Germany. *Regional Studies*, 40(9), 991–1004. <https://doi.org/10/ffxv8b>.
- CZSO. (2018). *Registr ekonomických subjektů—Počty subjektů podle okresu a CZ-NACE k 31. 5. 2018*. Czech Statistical Office.
- CZSO. (2019). *Statistical Yearbook of the Czech Republic 2019*. Czech Statistical Office. <https://www.czso.cz/csu/czso/statisticka-rocenka-ceske-republiky-2019>
- Delgado, M., Porter, M. E., & Stern, S. (2014). Clusters, convergence, and economic performance. *Research Policy*, 43(10), 1785–1799. <https://doi.org/10/f6nxzb>.
- Düzakin, E., & Düzakin, H. (2007). Measuring the performance of manufacturing firms with super slacks based model of data envelopment analysis: An application of 500 major industrial enterprises in Turkey. *European Journal of Operational Research*, 182(3), 1412–1432. <https://doi.org/10/b3k7v9>.
- Grant, J. L. (2003). *Foundations of economic value added* (2nd ed). Wiley.
- Koch, R., & Novotná, J. (2008). *Pravidlo 80/20: Umění dosáhnout co nejlepších výsledků s co nejmenším úsilím*. Management Press.
- Li, Z., Crook, J., & Andreeva, G. (2017). Dynamic prediction of financial distress using Malmquist DEA. *Expert Systems with Applications*, 80, 94–106. <https://doi.org/10/gg2bvn>.
- Lindqvist, G., Ketels, C., & Sölvell, Ö. (2012). *The cluster initiative greenbook 2.0*. Ivory Tower Publishers. <http://www.clusterobservatory.eu/system/modules/com.gridnine.opencms.modules.eco/providers/getpdf.jsp?uid=c57a2f9f-aa59-4af8-a8f9-4fa99e95b355>
- Lohr, S. L. (2010). *Sampling: Design and analysis* (2nd ed). Brooks/Cole.
- MPO. (2017). *Finanční analýza podnikové sféry za rok 2017*. Ministry of Industry and Trade. <https://www.mpo.cz/cz/rozcestnik/analyticke-materialy-a-statistiky/analyticke-materialy/financni-analyza-podnikove-sfery-za-rok-2017%2D%2D237570>
- Neumaierová, I., & Neumaier, I. (2002). *Výkonnost a tržní hodnota firmy*. Grada.
- Pantziros, C. J., Karagiannis, G., & Tzouvelekas, V. (2011). Parametric decomposition of the input-oriented Malmquist productivity index: With an application to Greek aquaculture. *Journal of Productivity Analysis*, 36(1), 21–31. <https://doi.org/10/fmv3g9>.
- Pavelková, D., & Jirčíková, E. (2008). Clusters as a tool to improve company's competitiveness. *E +M Ekonomie a Management*, 11(3), 62–75.
- Pavelková, D. (2013). *Klastrové politiky a jejich vliv na rozvoj klastrů a klastrových organizací*. Linde Praha.
- Pittnerová, R., & Rydvalová, P. (2014). *Networking pro výzkum a inovace*. VÚTS.
- Porter, M. E. (1998). Clusters versus Industrial Policy. In *On Competition*. Harvard Business School Press.
- Rydvalová, P., & Hotař, V. (2012). *Podnikání ve sklářství: Inovace jako cesta z krize s příklady ze severních Čech*. VÚTS.
- Rydvalová, P., & Pittnerová, R. (2013). Selection of potential partners from the Czech Glass industry for participation in EU framework programmes. *E+M Ekonomie a Management*, 16(1), 69–79.
- Rydvalova, P., & Zizka, M. (2021). Specifics of natural industry clusters. In M. Zizka & P. Rydvalova (Eds.), *Innovation and performance drivers of business clusters – An empirical study*. Springer Nature.
- Skokan, K. (2004). *Konkurenceschopnost, inovace a klastry v regionálním rozvoji*. Repronis.
- Skokan, K., Polednikova, E., & Stanickova, M. (2012). Establishment and growth of business clusters with public aid. *Journal of Competitiveness*, 4(3), 14–28. <https://doi.org/10/gg9ckb>.
- Stewart, G. B. (1994). EVA_{tm}: fast and fantasy. *Journal of Applied Corporate Finance*, 7(2), 71–84. <https://doi.org/10/b4fpr6>.
- Štrach, P. (2007). Writing teaching and research case studies. *Acta Oeconomica Pragensia*, 15(3), 22–36. <https://doi.org/10.18267/j.aop.62>.
- Svobodová, H., Věžník, A., & Hofmann, E. (2013). *Vybrané kapitoly ze socioekonomické geografie České republiky*. Masarykova univerzita. <http://is.muni.cz/elportal/?id=1104084>

- Wang, D. D. (2019). Performance assessment of major global cities by DEA and Malmquist index analysis. *Computers, Environment and Urban Systems*, 77, 101365. <https://doi.org/10/gg2bnr>.
- Žižka, M. (2006). Netzwerkstrukturen in der nordböhmisches und sächsischen Industrie. *Arbeit, Organisation Und Personal Im Transformationsprozess*, 11(25), 173–179.
- Žižka, M. (2013). *Hospodářský rozvoj regionů: Vymezení funkčních regionů, významné socioekonomické faktory, regionální odolnost a inovační intenzita*. Professional Publishing.
- Žižka, M., & Rydvalová, P. (2014). Influence of clusters on the intensity of innovation outputs. *Amfiteatru Economic*, 16(37), 994–1012.