

Article

Spatial Conditions Supporting Sustainable Development of Enterprises on Local Level

Aneta Ptak-Chmielewska *  and Agnieszka Chłoń-Domińczak 

Institute of Statistics and Demography, Warsaw School of Economics, 02-554 Warszawa, Poland; achlon@sgh.waw.pl

* Correspondence: aptak@sgh.waw.pl

Abstract: Micro, small and medium enterprises (MSMEs) represent more than 99% of enterprises in Europe. Therefore, knowledge about this sector, also in the spatial context is important to understand the patterns of economic and social development. The main goal of this article is an analysis of spatial conditions and the situation of MSMEs on a local level using combined sources of information. This includes data collected in the Social Insurance Institution and Tax registers in Poland, which provides information on the employment, wages, revenues and taxes paid by the MSMEs on a local level as well as contextual statistical information. The data is used for a diagnosis of spatial circumstances and discussion of conditions influencing the status of the MSMEs sector in a selected region (voivodeship) in Poland. Taxonomy methods including factor analysis and clustering methods based on k-means and SOM Kohonen were used for selecting significant information and grouping of the local units according to the situation of the MSMEs. There are eight factors revealed in principal component analysis and five clusters of local units distinguished using these factors. These include two clusters with a high share of rural local units and two clusters with a high share of rural-urban and urban local units. Additionally, there was an outstanding cluster with only two dominant urban local units. Factors show differences between clusters in the situation of MSMEs sector and infrastructure. Different spatial conditions in different regions influence the situation of MSMEs.

Keywords: micro, small and medium enterprises MSMEs; public registers; creating and developing MSMEs; data taxonomy methods; SOM Kohonen



Citation: Ptak-Chmielewska, A.; Chłoń-Domińczak, A. Spatial Conditions Supporting Sustainable Development of Enterprises on Local Level. *Sustainability* **2021**, *13*, 2292. <https://doi.org/10.3390/su13042292>

Academic Editor: Danuta Zawadzka

Received: 13 January 2021

Accepted: 17 February 2021

Published: 20 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Micro, small and medium enterprises (MSMEs) play a significant role in the economy. This sector represents more than 99% of all enterprises in Europe and they contribute significantly to local economies. Therefore, it is important to monitor the role of the MSMEs not only at the national or regional, but also at the local level. The MSMEs play various roles in the local economies. They create jobs and invest, contributing to the development of local economies, their infrastructure and growth potential. The MSMEs are also important actors that can contribute to sustainable development, as they are embedded in the local context and environment.

The relationships between the status of the MSMEs sector and the economic development at the regional level or local level recognized in the literature. Researchers focus both on the regional and local determinants of the development of this sector as well as the impact that MSMEs have on the economic development, competitiveness and diffusion of innovations at the local level.

The regional conditions shape the situation of the MSMEs. Literature findings indicate that the regional differences in industry density, income growth, population size growth, presence of the skilled labor force, accessibility of suppliers and customers and proximity of high quality universities have an impact on the enterprises development, including formation of new firms and their innovativeness as well as survival capacity [1–7]. Multiple

factors affect the creation process of new firms, their survival, and their innovative capacity both at the regional and local levels. From a literature review those factors include: population size and growth in the region, a high percentage share of people with high managerial and vocational skills in the population, urbanization level, households' wealth, and local demand from customers [6]. More precisely, local labor market conditions and potential involvement of the local labor force have an impact on the situation of enterprises [8–11]. The geographical concentration of industrial activities can generate agglomeration circumstances for start-ups and innovation as well as increases the level of competition and makes the exits of firms increase [12]. The research of such factors contributes to understanding the characteristics of the MSME sector situation and development in the regional context and helps to develop recommendations for local activities stimulating further development of the enterprises, including their innovativeness.

Some of the determinants of enterprise development depend on the spatial characteristics. These include agglomeration and urbanization level. Clustering of MSMEs can help identifying factors and local conditions that create favorable environment for MSMEs development [2,13]. Furthermore [14] it underlines that spatial clustering highlights the impact of proximity and distance, as well as institutions, and local distances on economic processes. Spatial clusters of enterprises increase diversity, observability, and comparability, and at the same time increase differentiation without lowering knowledge sharing. The spatial attributes of the interactive learning and innovation process could be a starting point not only for spatial agglomerations analysis, but also for developing research in the area economic geography.

Clusters in geographical space are quite natural due to distances between organizations, enterprises, and people. On one side, a short distance can stimulate interactions, learning, knowledge and experience sharing. On the other side, too small a space also increases competitiveness and lack of space for development. This can also apply to micro and small companies, which have not sufficient resources to reach out beyond the local context of activity. Understanding the diverse patterns of spatial distances at the local level is an important factor contributing to the identification of possible policies and solutions in developing innovation. In particular, urbanization plays a big role in local economies. It creates cooperation between enterprises, and makes the cost of transportation both for customers and suppliers lower [4]. The degree of urbanization also shapes the customers' demand. A higher urbanization level means both a higher density of customers and better access to the infrastructures. For example, access to transportation infrastructure has been found to have a significant effect on urban growth [15]. Finally, authors [16] also shows that regulations, including local fiscal policies and land use, are important factors explaining the urban sprawl, which is also linked to the density of the MSME sector.

Given the local character of many MSMEs, author [17] investigates the structure of such enterprises at the most local level. Those analyses provide more insight into variation in enterprise structure. Existing structure and distribution of MSMEs in spatial and economic circumstances gives information about factors determining such structure and distribution. Understanding of the current situation can also contribute to the planning of further development of this sector. According to the cited research, the impact of regional policy is dominating. Namely, the region becomes an entity on its own with assets (tangible and non-tangible) based on knowledge and institutional base. Those assets are accessible for local enterprises more than those available at the national level. However, even in the regional context, the spatial distribution of MSMEs is diversified. There is a clustering of economic activities in the sector, particularly in the localities that are more conducive for the development of entrepreneurship. The locally embedded values and attitudes toward entrepreneurship may also influence both the local rate and level of entrepreneurship [18,19].

All those factors determining MSMEs development are very much diversified among regions but also within regions. There is some evidence on that from Finland and the United Kingdom. In Finland [6], urban and rural communities give different environments

for enterprise development conditions, specifically with regard to human capital, access to the technology, the number of clusters of enterprises, and the intensity of cooperation in communities. In the UK [20], business growth is possible but under different territorial conditions, different levels of competition and demand between regions. The UK is also characterized by differences in the structure of occupational skills of the labor market.

The interrelations between local situation and the enterprises are in both directions. The MSME sector activities also stimulate economic growth, both nationally and regionally [21–27].

In Central and Eastern Europe, the experiences of the centrally planned economy and the economic transition from the 1990s shaped the development of the MSMEs. Author [28] shows that there is a regional divergence of the enterprise development in these countries, driven by the exogenous factors (i.e., distance to airports, borders with more developed countries or the transport corridors), endogenous factors (access to human capital, innovative businesses or advanced business services) as well as structural factors (rural regions or regions with heavy industry). The 2008 financial and economic crisis increased the regional polarization [29]. Authors [30] also underline that the CEE countries still lag behind Western Europe and the regional and local differences persist.

The determinants of the MSMEs sector development are also researched in Poland. According to [31], human capital development, wages levels, unemployment level, economic activity of the population in the region, and disposable income are crucial determinants of the development of the MSMEs sector. Analyses presented by [32] show that access to financial instruments at the regional level is one of the important barriers for the enterprise development. Factors stimulating such development are, among others, access to IT infrastructure, distance to market and suppliers, degree of cooperation between social partners.

Author [33] also warns that in particular in CEE, there is a risk that spatial development further concentrates in a smaller number of (metropolitan) regions, whereas more and more other regions might be affected by processes of peripheralization, which is against the idea of sustainable development.

Given these risks, there is a gap in the current literature, particularly in the inclusion of the spatial circumstances such as existing infrastructure, diversity of enterprises, and broader range of local factors, such as the financial situation of local governments and existing social infrastructure determining the MSMEs situation on a local level, which can help to inform policies that can reduce inequalities in the MSMEs development. The current literature usually focuses on regional differences. We claim that local factors have a significant impact on the situation of the MSMEs sector and differences within regions are substantial. This article contributes to filling this gap by adding spatial analysis based on clustering methods for a selected region at the level of the lowest local unit (gmina).

Our main goal is to analyze the spatial conditions supporting sustainable development of the MSMEs on a local level using obtained data from public registers (social insurance and tax authorities) at the local level as well as spatial data from Statistics Poland.

Our analysis focuses on the kujawsko-pomorskie voivodeship, which is one of the 16 regions in Poland. Kujawsko-pomorskie region was selected due to availability of local data obtained from administrative registers, but also as a region that has a similar economic situation to the entire Polish economy. We focus on the MSME sector, whose role in the Polish economy is substantial and increasing. In 2016, the MSMEs were responsible for the 49.8% of the Polish GDP, compared to 47.2% in 2008. This is mainly due to the increasing role of the small and medium enterprises, as the input of micro enterprises to GDP slightly decreased from 31.0% to 30.2%. The MSMEs contribution to the labor market demand is also significant. The share of employment in the MSME sector in total employment in Poland is stable on the level of 57% [34].

In the article, we focus on the identification of local differences in the situation of the MSME sector in the context of the overall economic and social situation. We use the taxonomy methods including principal components and factor analysis and clustering

methods based on k-means and SOM Kohonen. We propose these methods due to the simplicity of application and clear interpretation of results. We wanted to avoid “black-box” methods due to unclear rationale and interpretation.

The main goals of this article are as follows:

- Identification of the main factors affecting the situation of micro and small enterprises on local (gmina) level at the regional level.
- Diagnosis of regional differences of micro and small enterprises situation on local (gmina) level based in the kujawsko-pomorskie voivodship.

The analysis is based on the existing differences in the status of the MSMEs at the local level. While we focus on the one snapshot of time (2018 year), the analysis contributes to the identification of local factors conducive to the development of the MSMEs. It provides evidence that can be used for introducing local policies for supporting sustainable MSMEs, such as micro-lending, crowd funding, tokenization and community initiatives.

The article consists of four sections. After the brief introduction and literature review that was presented, we describe the materials and methods used in the analysis. Next we present our results and then conclude with discussion.

2. Materials and Methods

The analysis in the article is based on the three main data sources. The first one is data from an administrative register of the social Insurance Institution. It includes information on enterprises—payers of the social insurance contributions collected in the kujawsko-pomorskie voivodship. Second, we use data from the tax authorities on the revenues and taxes of paid by the enterprises that conduct their businesses in the local units in the region. This data was collected under the project REGIOGMINA (Usytuowanie na poziomie samorządów lokalnych instrumentów wsparcia dla Małych i Średnich Przedsiębiorstw (MŚP), działających w oparciu o model wielopoziomowego zarządzania regionem). The project is implemented by a consortium led by the regional government of the kujawsko-pomorskie voivodship with the SGH Warsaw School of Economics and Nicolaus Copernicus University in Torun, financed from the National Centre for Research and Development. We also used context data from the Statistics Poland Local Data Bank and its goal is to propose instruments that can support the MSMEs development in the region.

According to the Social Insurance Institution register of contribution payers in December 2018 there were 93,116 contribution payers in kujawsko-pomorskie voivodship (which is an estimate of the number of enterprises), of which 99.68% were:

- self-employed—51.29%;
- micro companies (up to 10 workers covered by social insurance)—38.3%;
- small companies (10–49 workers)—8.13%;
- medium companies (50–249 workers)—1.91%.

The largest share of the MSMEs were recorded in the trade (22%), construction (13%) and industrial processing (10%).

The main unit of the analysis is the local unit. There are 144 local units (gmina) in the region, which are included in the data. The distribution of enterprises by local units (gmina) is not univariate. In the five biggest cities in this voivodship: Bydgoszcz, Toruń, Włocławek, Grudziądz and Inowrocław there are 50.0% of self-employed activities, 47.0% of micro enterprises, 44.7% of small enterprises and 50.2% of medium enterprises located with the remaining 139 local units where the remaining share of the enterprises is located.

Variables used in this analysis were included in Table 1. These include two main types of variables. First is the data characterizing the MSMEs sector, including the structure of companies and employment by industry section, wage levels, employees of the MSMEs, revenues of the companies, the characteristics of local units, including the infrastructure, public finance and social infrastructure (proxied by the availability of preschools).

Table 1. Variables used in analysis.

Variable	Description	Source
service_share_payer	Share of contribution payers in Services	Social Insurance Institution
industry_share_payer	Share of contribution payers in Industry	Social Insurance Institution
avg_wage_f	Average contribution base—females	Social Insurance Institution
avg_wage_m	Average contribution base—males	Social Insurance Institution
avg_wage_micro	Average contribution base—micro enterprises	Social Insurance Institution
avg_wage_small	Average contribution base—small enterprises	Social Insurance Institution
avg_wage_medium	Average contribution base—medium enterprises	Social Insurance Institution
rev_pc_self	Average revenues of self-employed (taxes)	Tax administration
rev_pc_micro	Average revenues of micro enterprises (taxes)	Tax administration
rev_pc_small	Average revenues of small enterprises (taxes)	Tax administration
rev_pc_medium	Average revenues of medium enterprises (taxes)	Tax administration
ind_share_empl	Share of employed in Industry	Social Insurance Institution
serv_share_empl	Share of employed in Services	Social Insurance Institution
micro_per10ths	Number of micro-enterprises per 10ths citizens	Social Insurance Institution
small_per10ths	Number of small enterprises per 10ths citizens	Social Insurance Institution
medium_per10ths	Number of medium enterprises per 10ths citizens	Social Insurance Institution
type_region	Type of Region	Statistics Poland
roads_regional_km ²	Regional roads (km) per squared km	Statistics Poland
roads_voivodeship_km ²	Voivodeship roads (km) per squared km	Statistics Poland
roads_local_km ²	Local roads (km) per squared km	Statistics Poland
perc_ground_roads	Percentage of ground roads	Statistics Poland
pre_school_education	Number of children 3–5 age in pre-school education per 1000 of 3–5 age children	Statistics Poland
expenses_pc	Local government expenditure per capita	Statistics Poland
expenses_educ_pc	Local government expenditure for education per capita	Statistics Poland
expenses_capital_pc	Local government capital expenditure per capita	Statistics Poland
Feminization	Feminization ratio (number of women per 100 men)	Statistics Poland
pop_prod_share	Share of population in working age	Statistics Poland
unemp_prod_share	Unemployment rate	Statistics Poland
rev_pc	Revenues per capita	Statistics Poland
rev_form_taxes_pc	Revenues from personal income taxes per capita	Statistics Poland
debt_pc	Debt per capita of the local government	Ministry of Finance

Source: Authors' analysis.

Two types of taxonomy methods were used in this analysis to verify the research hypothesis. The first method is two-stage. We performed factor analysis to identify factors determining the MSMEs situation in the local units from the available database. We applied the principal component method and quartimax rotation, for better interpretation of results and identification of implicit factors that contribute to the differentiation of the MSMEs status at the local level.

The factor analysis was based on principal component method to find factor weights. Using eigenvalues and eigenvectors, linear combinations of variables were calculated with coefficients driven by eigenvectors components. Those linear combinations give the highest possible proportion of variance explained. The first few combinations are used for factors with the highest eigenvalue (highest proportion of explained variance). Orthogonal rotation (for example quartimax) is applied to give better understanding and interpretability of results. High coefficient (factor weight) is a high correlation between variable and factor. Final combination can be interpreted as hidden factor based on factor weights.

The second method was the clustering. We used a K-means method as the most frequently used method that was applied. Euclidean distance is used as the default distance measure. The number of clusters is determined for the start and the next randomly clustered seeds are chosen. Each observation ($i = 1, \dots, n$) is classified to the group with nearest cluster seed measured by Euclidean distance. For all clusters ($j = 1, \dots, k$), new cluster centers are calculated as arithmetic mean, all observations belonging to the group. Those steps are repeated until there are no other moves between groups. Error function

is calculated at each step—sum of quadratic distance intergroup calculated from groups centers:

$$F = \sum_{j=1}^k \sum_{O_i \in S_j} d(O_i, M_j)^2 \quad (1)$$

where d is Euclidean distance.

In practice this process is convergent after a few iterations, but in general as this algorithm does not have to be convergent the maximum number of iterations is pre-defined.

To compare the results of k-means algorithm the algorithm based on unsupervised learning was applied. Self-Organizing Maps by Kohonen (SOM) is a basic method of unsupervised learning proposed and optimized by [35]. This algorithm reflects variables values into two-dimensional space. The number of records must be higher than number of cells in this network (so higher than number of proposed clusters). Each observation gets its representation in the network (neurons) by appropriate mapping. Neurons with highest frequency of observations tighten the placement around the neuron and the process is continued for all remaining observations. This algorithm generates a network with neurons (clusters) reflecting the distance between variables into new dimension. The basic difference between k-means method and SOM Kohonen network is that in SOM Kohonen the distances between clusters are based on the network and not on the original space. Results are heavily dependent on different measurement scales of variables, which is why variables must be standardized.

3. Results

The factor analysis revealed hidden factors showing diversity of the situations of MSMEs in different local units (gmina). Based on eigenvalue criterion (above 1) and reasonable share of explained variance, eight factors were selected (see Table 2).

Table 2. Factor analysis results—quartimax rotation.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
pop_prod_share	−0.20	−0.41	0.19	0.32	0.12	0.26	−0.10	−0.12
unemp_prod_share	−0.30	−0.17	0.27	−0.52	−0.11	−0.11	0.35	0.30
rev_pc	0.72	0.32	0.00	0.24	0.25	0.03	0.10	−0.05
rev_from_taxes_pc	0.74	0.31	−0.08	0.39	0.15	0.04	0.11	−0.02
debt_pc	0.03	−0.03	0.08	0.09	0.65	−0.08	0.24	−0.19
service_share_payer	0.29	0.17	−0.12	0.13	−0.02	0.04	0.84	0.00
industry_share_payer	−0.30	−0.22	0.09	0.01	0.00	−0.05	−0.79	−0.01
avg_wage_f	0.20	0.07	−0.05	0.58	−0.03	−0.41	0.02	0.49
avg_wage_m	0.24	0.06	0.09	0.78	0.08	−0.12	0.08	0.30
avg_wage_micro	0.01	−0.09	0.11	0.16	0.15	0.10	0.01	0.73
avg_wage_small	0.34	0.11	−0.06	0.65	0.10	−0.01	0.12	0.24
avg_wage_medium	0.11	0.07	0.14	0.75	0.07	−0.01	0.01	−0.13
rev_pc_self	0.25	0.92	−0.04	0.14	0.09	0.01	0.08	−0.03
rev_pc_micro	0.19	0.94	−0.06	0.09	0.04	0.00	0.10	−0.03
rev_pc_small	0.20	0.95	−0.06	0.09	0.07	0.01	0.07	−0.04
rev_pc_medium	0.09	0.94	−0.03	0.03	0.05	0.01	0.04	−0.02
ind_share_empl	0.04	−0.06	−0.04	−0.07	−0.05	0.94	−0.05	0.06
serv_share_empl	0.10	0.04	−0.03	−0.05	−0.02	0.95	0.12	0.02
micro_per10ths	0.81	0.12	−0.07	0.26	0.02	0.12	0.16	−0.18
small_per10ths	0.79	−0.01	−0.08	0.17	0.03	0.20	−0.13	−0.11
medium_per10ths	0.80	0.10	−0.01	−0.18	−0.06	−0.16	0.13	0.15
roads_regional_km ²	−0.14	−0.08	0.60	−0.17	0.09	−0.12	−0.29	0.10
roads_voivodeship_km ²	−0.11	−0.04	0.71	0.09	−0.03	0.03	0.10	−0.15
roads_local_km ²	−0.05	−0.06	0.84	0.11	0.15	0.03	−0.02	0.10
perc_ground_roads	−0.11	−0.08	0.95	0.02	0.14	−0.03	−0.11	0.09
pre_school_education	0.60	0.19	−0.21	0.12	−0.10	−0.04	0.18	0.24
expenses_pc	0.06	0.14	0.22	−0.02	0.89	−0.01	−0.09	0.09
expenses_educ_pc	−0.01	0.13	−0.17	0.15	0.59	0.07	−0.14	0.15
expenses_capital_pc	0.05	−0.03	0.22	0.05	0.83	−0.04	−0.05	0.10
feminization	0.55	0.39	−0.17	−0.09	−0.14	−0.10	0.45	0.13

Source: Authors' analysis.

Factors influencing the situation of micro, small and medium enterprises on local level:

- Factor 1—access to pre-school education, feminization ratio, number of micro, small and medium enterprises.
- Factor 2—average revenues of micro, small and medium enterprises.
- Factor 3—roads infrastructure: voivodehip, regional, local and ground roads.
- Factor 4—average contribution base from small, medium enterprises, average contribution base of females and males (negative correlation with unemployment).
- Factor 5—public debt of local unit (gmina), expenditure of local units (gmina) total, expenditure for education and capital expenditure per capita.
- Factor 6—share of employed in industry and services.
- Factor 7—share of contribution payers in services (negative correlation with share of contribution payers in industry)
- Factor 8—average contribution base of women and in micro enterprises—females, average wage based on contribution—micro enterprises.

Among these factors there are two that are related directly to the environment in which the MSMEs operate: Factor 5 that reflects the situation of the finances of the local unit and Factor 3 related to the infrastructure. Factor 1 reflects both the density of the sector and the local social conditions, while the remaining factors reflect the financial situation of the sectors and employment characteristics.

Hierarchical cluster analysis based on minimal Ward variance method was used as a pre-step for k-means clustering. To find an optimal number of clusters pseudo t-sqrt statistic was utilized because Cubic Clustering Criterium (CCC) and pseudo-F were much less informative. According to this measure the optimal number of clusters was set up as five clusters (see Figure 1).

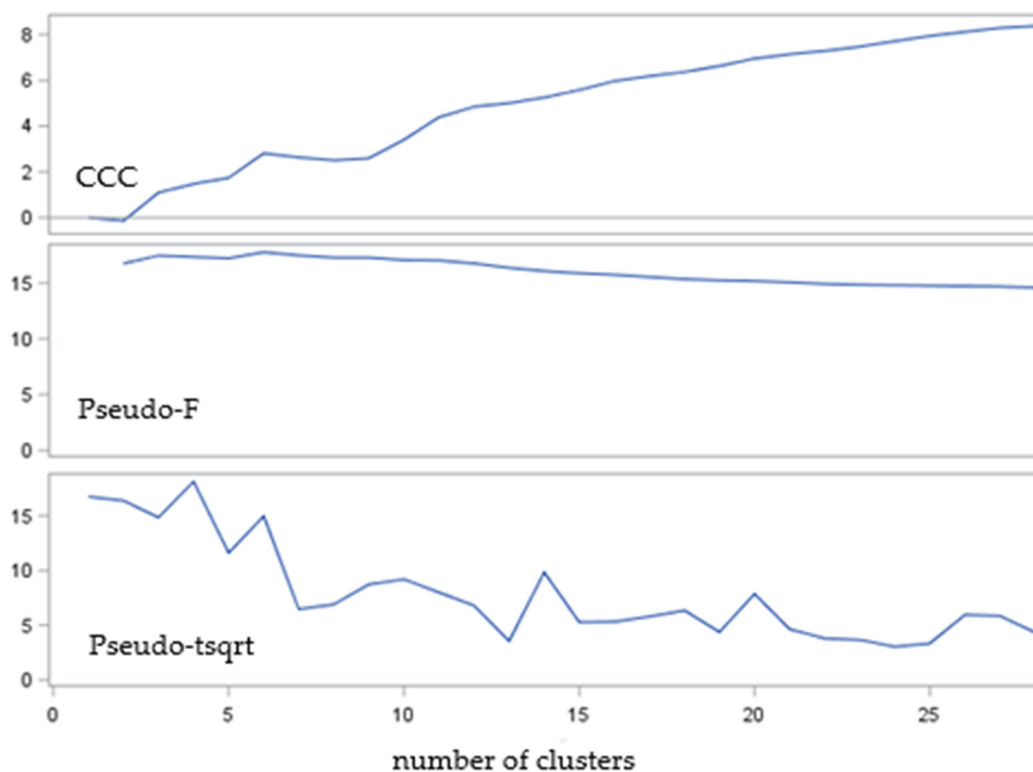


Figure 1. Hierarchical clustering—optimal number of clusters. Source: Authors' analysis.

3.1. K-Means Clustering

As a result of the clustering, two clusters with a high share of rural local units (gmina) emerged (Cluster 1 and Cluster 2), all together 25 local units and two clusters with high share of rural-urban and urban local units (gmina) (Cluster 3 and Cluster 4), all together 117 local units. Additionally, there was an outstanding cluster 5 with only two urban local units: Toruń and Budgoszcz (see Table 3).

Table 3. Local units (gmina) structure—k-means clusters.

Local Unit Type	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Total
urban	0 (0.0%)	0 (0.0%)	12 (14.5%)	3 (8.8%)	2 (100.0%)	17 (11.8%)
urban-rural	1 (8.3%)	0 (0.0%)	24 (28.9%)	10 (29.4%)	0 (0.0%)	35 (24.3%)
rural	11 (91.7%)	13 (100.0%)	47 (56.6%)	21 (61.8%)	0 (0.0%)	92 (63.9%)
Total	12 (100.0%)	13 (100.0%)	83 (100.0%)	34 (100.0%)	2 (100.0%)	144 (100.0%)

Source: Authors' analysis.

Spatial differences on the local level are visible on the local units (gmina) type differences: urban, rural and mixed (see Table 3). Results of k-means clustering are the following (see Tables 4 and 5 and Figure 2):

- (n = 12) Cluster 1—cluster is characterized by a high number of micro and small enterprises per 10ths citizens (see Table 3). There is majority of rural local units (gmina) and the most important are Factor 1, 4 and 5 (see Figure 2). This cluster is characterized by high average wage, high feminization ratio and high pre-school education ratio. There is also low unemployment and medium level of expenses in local units (see Table 4). These local units surround the two largest cities in the region (cluster 5).
- (n = 13) Cluster 2—cluster is characterized by low number of micro and small enterprises per 10ths citizens (see Table 3). There are only rural local units (gmina) in this cluster and the most important is Factor 6 (see Figure 2) characterized by a high share employed in industry and services, very low average wage, high unemployment rate and weak road infrastructure (see Table 4).
- (n = 83 + 34) Clusters 3 and 4—those clusters are characterized by the lowest number of micro and small enterprises per 10ths citizens (see Table 3). Those two clusters are typical medium profile clusters with the highest frequency of local units (gmina) and smallest differences between clusters. In cluster 3 there are weaker roads infrastructure. In cluster 4 there is good roads infrastructure but higher expenses per capita on local unit level (gmina). There is medium average wage and a low share working in industry and services in those two clusters (see Table 4).
- (n = 2) Cluster 5—this cluster is outstanding with only two local units, only two urban units with two main cities in this region (see Table 3): Toruń and Bydgoszcz. There is a high number of enterprises in this region. This cluster is also characterized by low unemployment and relatively high average wage. There are very high revenues from micro and small enterprises. There is a very high feminization ratio and pre-school education share in those two main cities (see Table 5).

Table 4. Number of enterprises per 10ths citizens—k-means clusters.

Variable	Mean	Std. Dev.	Min	Max	N	Median
Cluster 1						
micro_per10ths	201.53	47.92	136.39	307.01	12	189.39
small_per10ths	45.71	12.78	24.49	68.85	12	45.17
medium_per10ths	8.30	3.27	4.08	12.97	12	8.99
Cluster 2						
micro_per10ths	145.00	34.02	84.19	212.03	13	142.30
small_per10ths	34.47	12.45	18.94	52.44	13	33.39
medium_per10ths	3.14	3.44	0	10.20	13	2.39
Cluster 3						
micro_per10ths	139.80	38.57	25.60	240.65	83	134.85
small_per10ths	32.04	10.48	8.15	65.73	83	31.14
medium_per10ths	6.24	3.84	0	19.82	83	5.95
Cluster 4						
micro_per10ths	131.79	36.61	61.73	251.53	34	124.42
small_per10ths	29.20	11.87	8.81	66.05	34	27.06
medium_per10ths	7.16	3.57	2.17	17.85	34	6.35
Cluster 5						
micro_per10ths	211.79	18.07	199.01	224.57	2	211.79
small_per10ths	42.17	2.78	40.21	44.14	2	42.10
medium_per10ths	10.90	1.24	10.02	11.78	2	10.90

Source: Authors' analysis.

Table 5. K-means analysis—mean values for clusters (5 clusters).

Variable	Cluster 1 (12)	Cluster 2 (13)	Cluster 3 (83)	Cluster 4 (34)	Cluster 5 (2)
pop_prod_share	0.630	0.627	0.620	0.622	0.592
unemp_prod_share	0.036	0.065	0.069	0.085	0.033
rev_pc	2721.990	1631.760	1632.180	1784.680	3332.070
rev_from_taxes_pc	1128.850	560.988	587.809	549.584	1413.040
debt_pc	1477.370	1023.750	922.092	1441.160	1021.520
service_share_payer	53.216	49.758	48.017	48.556	60.413
industry_share_payer	27.974	28.985	31.001	30.048	18.642
avg_wage_f	2894.500	2222.240	2609.300	2556.820	2782.710
avg_wage_m	3219.620	2731.370	2886.480	2894.490	3098.540
avg_wage_micro	2206.020	2117.830	2103.260	2142.750	2056.860
avg_wage_small	3178.570	2675.860	2683.310	2648.370	3089.120
avg_wage_medium	3659.490	3353.190	3350.000	3332.190	3571.580
rev_pc_self	328.37 mln	65.95 mln	100.79 mln	111.32 mln	4095.29 mln
rev_pc_micro	327.69 mln	95.59 mln	164.79 mln	169.66 mln	5531.95 mln
rev_pc_small	313.40 mln	105.14 mln	134.12 mln	155.61 mln	4766.88 mln
rev_pc_medium	246.93 mln	24.86 mln	141.23 mln	237.81 mln	12,159.49 mln
ind_share_empl	0.190	0.571	0.211	0.198	0.169
serv_share_empl	0.261	0.702	0.212	0.216	0.341
micro_per10ths	201.533	145.003	139.801	131.788	211.792
small_per10ths	45.709	34.469	32.046	29.201	42.175
medium_per10ths	8.299	3.140	6.241	7.165	10.901
roads_regional_km ²	0.043	0.075	0.100	0.229	0.000
roads_voivodeship_km ²	0.007	0.022	0.013	0.064	0.008
roads_local_km ²	0.214	0.162	0.075	0.471	0.093
perc_ground_roads	26.372	25.845	18.747	76.322	10.220
pre_school_education	777.083	595.154	690.602	650.971	942.000
expenses_pc	5636.510	5150.210	4805.670	5675.970	6054.870
expenses_educ_pc	1693.160	1435.700	1377.120	1509.300	1729.010
expenses_capital_pc	1476.500	1030.230	793.301	1387.470	1242.000
feminization	102.083	98.923	101.434	100.882	114.000

Source: Authors' analysis.

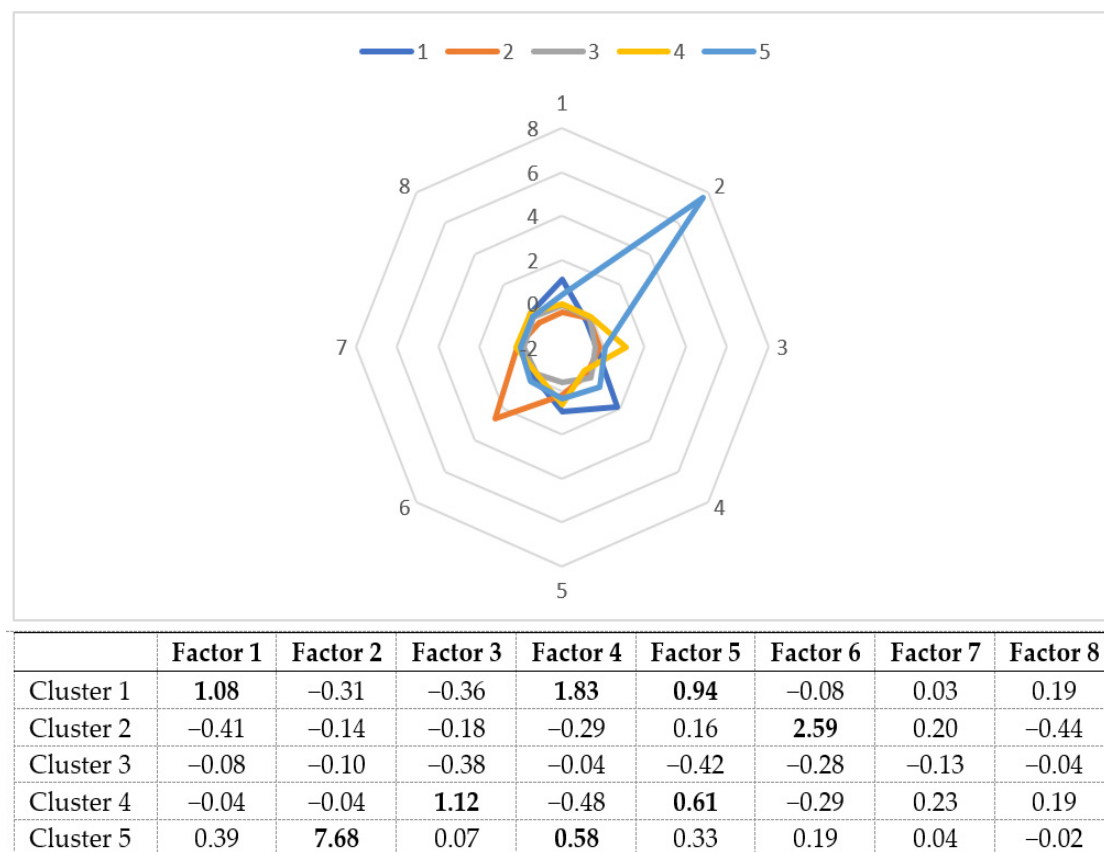


Figure 2. Factors' averages for k-means clustering. Source: Authors' analysis.

3.2. SOM Kohonen Clustering

Due to an even number of dimensions in SOM Kohonen the outstanding cluster 5 was excluded and grouping into four clusters was applied on the remaining sample of local units (gmina). SOM Kohonen was performed on eight factors.

There are two clusters with high share of rural local units (Cluster 3 and Cluster 4) all together with 59 local units (gmina) and two clusters with relatively high share of urban-rural and urban local units (Cluster 1 and Cluster 2) all together with 83 local units (gmina) (see Table 6).

Table 6. Local units (gmina) structure—Self-Organizing Maps by Kohonen (SOM) Kohonen clusters.

Local Unit Type	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Total
urban	11 (18.3%)	2 (8.8%)	1 (2.9%)	1 (4.1%)	2 (100.0%)	17
urban-rural	18 (30.0%)	7 (30.4%)	6 (17.1%)	4 (16.7%)	0 (0.0%)	35
rural	31 (51.7%)	14 (60.8%)	28 (80.0%)	19 (79.2%)	0 (0.0%)	92
Total	60 (100.0%)	23 (100.0%)	35 (100.0%)	24 (100.0%)	2 (100.0%)	144

Source: Authors' analysis.

Spatial differences on a local level are visible on the local units (gmina) type differences: urban, rural and mixed. Results of SOM Kohonen clustering are the following:

- (60) Cluster 1—clusters with a significant role of Factor 4 and Factor 7, mostly correlated with average wage based on contribution, especially in the services sector (see Figure 3). In this cluster there is a high average number of micro enterprises and also a high number of medium enterprises (see Table 6). Finally, local units that belong to this cluster have a small share of ground roads, low capital expenses in the region, a

higher feminization ratio as well as quite a high coverage of the pre-school education (see Table 7).

- (23) Cluster 2—significant importance of Factor 3, 6 and 8 (see Figure 3). Factor 8 is strongly correlated with average wage (contribution base) of women and average wage in micro-enterprises. Cluster 2 is characterized by a low number of small and micro enterprises and slightly higher number of medium enterprises (see Table 7). At the same time, these local units have a very low level of revenues per capita but high level of expenses (see Table 8).
- (35) Cluster 3 and (24) Cluster 4—significant importance of Factor 3 and Factor 6 (see Figure 3). Factor 3 is correlated with roads infrastructure: voivodeship, regional, local and ground. Factor 6 is correlated with share of those employed in industry and services. Lowest number of small and medium enterprises. Those clusters are also characterized by lowest revenues per capita. Surprisingly, those clusters have lowest feminization ratio (see Table 8).
- (2) Cluster 5—already described in k-means section. Very high number of micro and small enterprises (see Table 7). In this outlying cluster there are only two main cities in the region: Toruń and Bydgoszcz. This cluster is characterized by very low unemployment, and high average wage based on contribution. There are very high revenues from micro and small enterprises. In those two main cities there is a very high feminization ratio and pre-school education share (see Table 8).

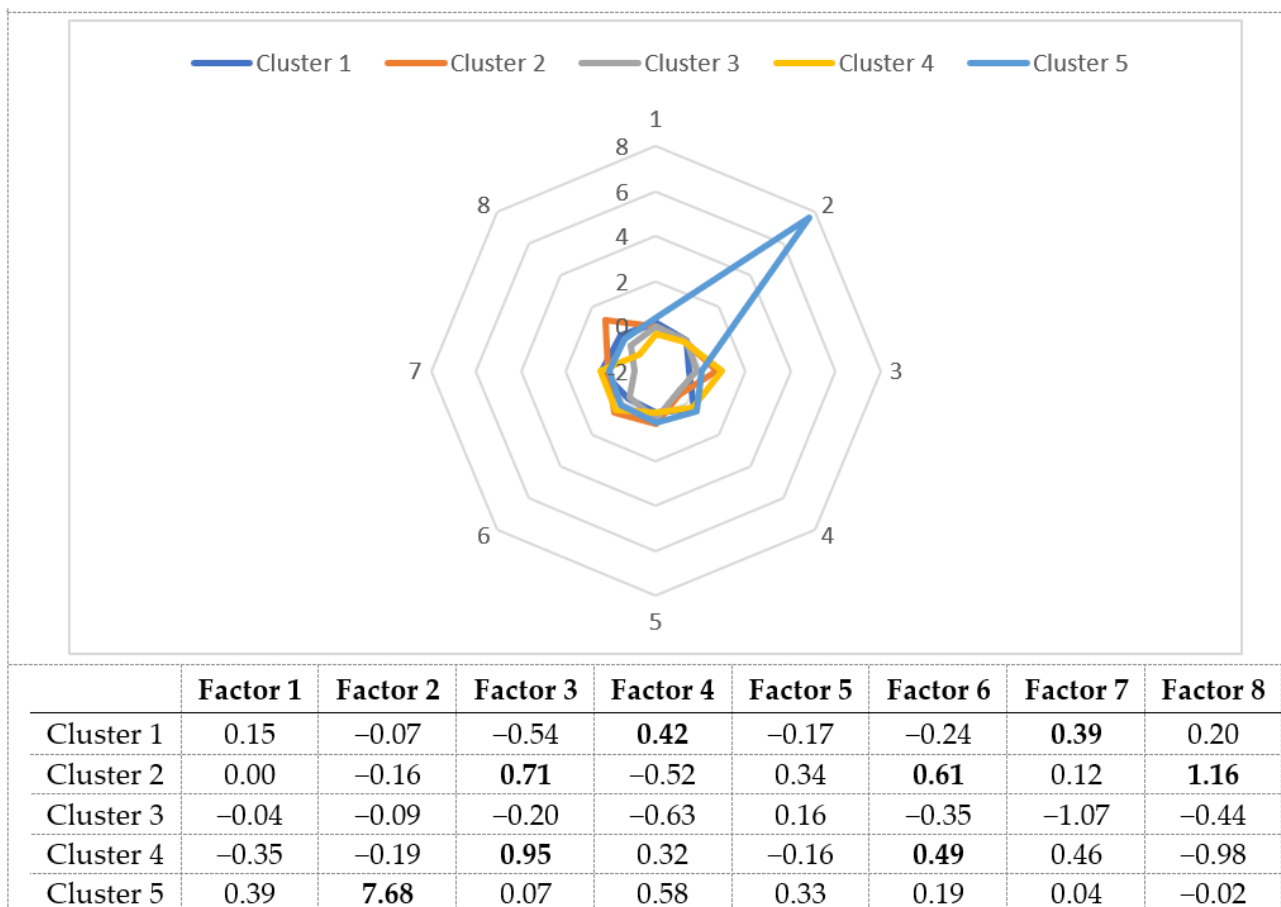


Figure 3. Factors' averages for SOM Kohonen clustering. Source: Authors' analysis.

Table 7. Number of enterprises per 10ths citizens—SOM Kohonen clusters.

Variable	Mean	Std. Dev.	Min	Max	N	Median
Cluster 1						
micro_per10ths	153.73	43.14	74.63	307.01	60	151.06
small_per10ths	34.15	11.42	11.76	68.85	60	34.23
medium_per10ths	7.12	3.40	0	13.81	60	6.84
Cluster 2						
micro_per10ths	132.57	40.02	61.73	231.53	23	125.61
small_per10ths	32.12	13.51	15.65	66.05	23	29.58
medium_per10ths	6.77	5.11	0	17.85	23	5.78
Cluster 3						
micro_per10ths	131.36	44.02	25.60	251.53	35	126.30
small_per10ths	32.61	11.56	8.15	63.47	35	30.78
medium_per10ths	5.53	3.71	0	19.82	35	4.74
Cluster 4						
micro_per10ths	146.54	33.79	93.95	231.49	24	145.42
small_per10ths	29.99	11.94	8.81	65.73	24	27.86
medium_per10ths	5.23	3.48	0	13.27	24	5.08
Cluster 5						
micro_per10ths	211.79	18.07	199.01	224.57	2	211.79
small_per10ths	42.17	2.78	40.21	44.14	2	42.17
medium_per10ths	10.90	1.24	10.02	11.78	2	10.90

Source: Authors' analysis.

Table 8. Mean values for SOM Kohonen clusters.

Variable	Cluster 1 (60)	Cluster 2 (23)	Cluster 3 (35)	Cluster 4 (24)	Cluster 5 (2)
pop_prod_share	0.62	0.62	0.62	0.63	0.59
unemp_prod_share	0.07	0.08	0.07	0.07	0.03
rev_pc	1905.42	1727.84	1613.87	1644.82	3332.07
rev_pit_pc	728.28	516.54	540.44	575.85	1413.04
debt_pc	1097.44	1124.46	956.56	1307.58	1021.52
service_share_payer	52.21	48.20	42.32	49.96	60.41
industry_share_payer	27.67	28.92	35.92	30.19	18.64
avg_wage_f	2704.48	2574.69	2489.08	2438.42	2782.71
avg_wage_m	2984.65	2910.44	2765.61	2888.25	3098.54
avg_wage_micro	2139.13	2266.89	2043.30	2059.45	2056.86
avg_wage_small	2816.66	2672.09	2582.54	2701.76	3089.12
avg_wage_medium	3426.75	3275.61	3272.36	3473.90	3571.58
rev_pc_self	197.86 mln	39.77 mln	66.01 mln	77.18 mln	4095.29 mln
rev_pc_micro	283.47 mln	77.28 mln	93.58 mln	106.68 mln	5531.96 mln
rev_pc_small	248.22 mln	57.52 mln	82.84 mln	101.47 mln	4766.88 mln
rev_pc_medium	289.84 mln	81.78 mln	79.27 mln	43.69 mln	12,159.49 mln
ind_share_empl	0.21	0.33	0.20	0.29	0.17
serv_share_empl	0.23	0.37	0.19	0.33	0.34
micro_per10ths	153.74	132.57	131.36	146.54	211.79
small_per10ths	34.15	32.12	32.61	29.99	42.18
medium_per10ths	7.12	6.77	5.53	5.24	10.90
roads_regional_km ²	0.06	0.20	0.17	0.12	0.00
roads_voivodeship_km ²	0.01	0.04	0.01	0.08	0.01
roads_local_km ²	0.06	0.40	0.13	0.39	0.09
perc_ground_roads	13.08	64.44	31.38	59.91	10.22
pre_school_education	746.55	695.48	634.89	562.71	942.00
expenses_pc	4924.51	5463.85	5260.60	5049.36	6054.87
expenses_educ_pc	1462.66	1517.07	1433.41	1324.04	1729.01
expenses_capital_pc	857.15	1259.39	1133.54	1002.50	1242.00
feminization	103.10	100.87	98.91	99.67	114.00

Source: Authors' analysis.

The analysis resulted in the creation of classifications based on the dominant characteristic of the local units assigned to the specific cluster. There is a clear division into rural-urban sub-regions with low infrastructure far from urban centers. In those sub-regions we have lowest number of enterprises per 10ths citizens. We also observe two rural sub-regions with a high number of enterprises surrounding two outlier urban regions (main cities in the voivodeship). The infrastructure around the biggest cities is better than in the peripheral areas. This means that the local units surrounding big cities benefit from the proximity of the local centers, including high average wage, high feminization ratio and high pre-school education ratio. They also have low unemployment and a medium level of expenditure of the local governments.

4. Discussion

Spatial analysis revealed differences between local units that form distinct clusters. There is also a visible dependence on the urbanization level of regions (see Figure 4). The two largest cities in the region: Bydgoszcz and Toruń sprawl their urbanization effect on the surrounding local units. Distinct local units are also concentrated around the two remaining big cities (Grudziądz, Włocławek) and are mostly urban-rural specific. At the same time, there is a large group of local units, who are weaker economically and more peripheral.

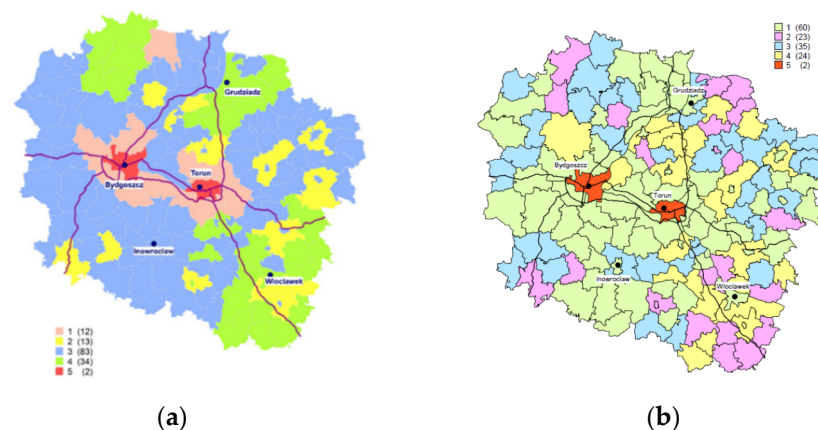


Figure 4. (a) K-means clusters map (b) SOM Kohonen clusters map.

Our analysis helps to identify features that differentiate the local units in the region, which include both factors related to the situation of the MSMEs sector and the environment in which they operate. The factor that explains the highest share of the observed variance covers both the characteristics of the sector (density of the MSMEs at the local level), but also demographic structure (feminization rate) and social infrastructure (pre-school coverage). This shows that the condition of the MSMEs sector is closely linked to the social infrastructure and demographic development. Such a combined “hidden” factor confirms that at the local level the conditions for sustainable development are linked to the role of the MSMEs sector.

The next identified factors include two that are related to the financial situation of the companies (revenues) and the wages they offer and two factors related to road infrastructure and the situation of public finance at the local level. This shows that the differences that are observed between groups of the local units stem both from the differences within the MSMEs sector, but also the environment they operate in, which are interconnected.

Based on those factors some specific policies can be addressed to support the MSMEs sector sustainable development on the local level, in particular designing local policies and instruments creating conditions conducive to its growth. In particular, we see that the situation in the local units surrounding the larger cities is relatively better, so there is an urbanization spread from which they benefit. As a result, there is little need of further

active public policies focusing on supporting the MSMEs development in these clusters. In contrast, the urban-rural areas more distant from the big cities have a lower average wage and low share of work in industry and services. At the same time, these local units have less financial potential to stimulate MSMEs development. Therefore, maintaining more harmonious development and reduction in the observed differences requires support from regional policy makers, including investment in infrastructure and instruments that support entrepreneurship (i.e., access to microfinance) that in turn lead to job creation, reduction in unemployment and sustainable development.

The main limitation of this research is that it is based on the observation in one year only in the area limited to one region. It can be extended in the future by the possible inclusion of dynamic information that could influence the results. Despite these deficiencies it shows that broader use of the administrative data can provide a deeper insight into the situation of the MSMEs at the local level. Future research will cover a wider geographic area and also on dynamic approach. In the authors' opinion, this could contribute to a better understanding of the local determinants of the MSMEs condition and the role of the MSMEs sector in sustainable development at the local level.

Author Contributions: Conceptualization, A.P.-C. and A.C.-D.; methodology, A.P.-C.; software, A.P.-C.; validation, A.P.-C. and A.C.-D.; formal analysis, A.P.-C.; investigation, A.P.-C.; resources, A.C.-D.; data curation, A.C.-D.; writing—original draft preparation, A.P.-C.; writing—review and editing, A.P.-C.; visualization, A.P.-C. and A.C.-D.; supervision, A.P.-C.; project administration, A.P.-C.; funding acquisition, A.C.-D. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for this research was provided by Warsaw School of Economics.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Armington, C.; Acs, Z.J. The determinants of regional variation in new firm formation. *Reg. Stud.* **2002**, *36*, 33–45. [[CrossRef](#)]
2. Boschma, R.A.; Wenting, R. The spatial evolution of the British automobile industry: Does location matter? *Ind. Corp. Chang.* **2007**, *16*, 213–238. [[CrossRef](#)]
3. Bruno, A.; Tyebjee, T. The environment for entrepreneurship. In *Encyclopedia of Entrepreneurship*; Kent, C.A., Sexton, D.L., Vesper, K.H., Eds.; Prentice-Hall: Englewood Cliffs, NJ, USA, 1982; pp. 288–307.
4. Elert, N. What determines entry? Evidence from Sweden. *Ann. Reg. Sci.* **2014**, *53*, 55–92. [[CrossRef](#)]
5. Lee, S.; Florida, R.; Acs, Z. Creativity and entrepreneurship: A regional analysis of new firm formation. *Reg. Stud.* **2004**, *38*, 879–891. [[CrossRef](#)]
6. Ritsila, J.J. Regional differences in environments for enterprises. *Entrep. Reg. Dev.* **1999**, *11*, 187–202. [[CrossRef](#)]
7. Van de Ven, A.H. The development of an infrastructure for entrepreneurship. *J. Bus. Ventur.* **1993**, *8*, 211–230. [[CrossRef](#)]
8. Helfat, C.E.; Lieberman, M.B. The birth of capabilities: Market entry and the importance of pre-history. *Ind. Corp. Chang.* **2002**, *11*, 725–760. [[CrossRef](#)]
9. Klepper, S. Employee startups in high-tech industries. *Ind. Corp. Chang.* **2001**, *10*, 639–674. [[CrossRef](#)]
10. Koo, J.; Cho, K. New firm formation and industry clusters: A case of the drugs industry in the US. *Growth Chang.* **2011**, *42*, 179–199. [[CrossRef](#)]
11. Shane, S. Prior knowledge and the discovery of entrepreneurial opportunities. *Organiz. Sci.* **2000**, *11*, 448–469. [[CrossRef](#)]
12. Boschma, R.A.; Frenken, K. Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *J. Econ. Geogr.* **2006**, *6*, 273–302. [[CrossRef](#)]
13. Guo, Q.; He, C.; Li, D. Entrepreneurship in China: The role of localisation and urbanisation economies. *Urban. Stud.* **2016**, *53*, 2584–2606. [[CrossRef](#)]
14. Malmberg, A.; Maskell, P. The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environ. Plan.* **2002**, *34*, 429–449. [[CrossRef](#)]
15. Duranton, G.; Turner, M.A. The Fundamental Law of Road Congestion: Evidence from US Cities. *Am. Econ. Rev.* **2011**, *101*, 2616–2652. [[CrossRef](#)]
16. Burchfield, M.; Overman, H.G.; Puga, D.; Turner, M.A. Causes of sprawl: A portrait from space. *Q. J. Econ.* **2006**, *121*, 587–633. [[CrossRef](#)]

17. Boschma, R.A. Competitiveness of regions from an evolutionary perspective. *Reg. Stud.* **2004**, *38*, 1001–1014. [[CrossRef](#)]
18. Westlund, H.; Bolton, R. Local social capital and entrepreneurship. *Small Bus. Econ.* **2003**, *21*, 77–113. [[CrossRef](#)]
19. Fritsch, M.; Wyrwich, M. The Long Persistence of Regional Entrepreneurship Culture: Germany 1925–2005. *Reg. Stud.* **2014**, *48*, 939–954. [[CrossRef](#)]
20. Vaessen, P.; Keeble, D. Growth-oriented SMEs in Unfavorable Regional Environments. *Reg. Stud.* **1995**, *29*, 489–505. [[CrossRef](#)]
21. Acemoglu, D.; Aghion, P.; Zilibotti, F. Distance to frontier, selection, and economic growth. *J. Eur Econ. Assoc.* **2006**, *4*, 37–74. [[CrossRef](#)]
22. Acemoglu, D.; Aghion, P.; Lelarge, C.; van Reenen, J.; Zilibotti, F. Technology, information, and the decentralization of the firm. *Q. J. Econ.* **2007**, *122*, 1759–1799. [[CrossRef](#)]
23. Aghion, P.; Bloom, N.; Blundell, R.; Griffith, R.; Howitt, P. Competition and innovation: An inverted-U relationship. *Q. J. Econ.* **2005**, *120*, 701–728.
24. Aghion, P.; Burgess, R.; Redding, S.; Zilibotti, F. Entry and productivity growth: Evidence from microlevel panel data. *J. Eur. Econ. Assoc.* **2004**, *2*, 265–276. [[CrossRef](#)]
25. Aghion, P.; Griffith, R. *Competition and Growth: Reconciling Theory and Evidence*; University College: London, UK, 2008.
26. Van Stel, A.; Carree, M.; Thurik, R. The effect of entrepreneurial activity on national economic growth. *Small Bus. Econ.* **2005**, *24*, 311–321. [[CrossRef](#)]
27. Wennekers, S.; Thurik, R. Linking entrepreneurship and economic growth. *Small Bus. Econ.* **1999**, *13*, 27–56. [[CrossRef](#)]
28. Smełkowski, M. The role of exogenous and endogenous factors in the growth of regions in Central and Eastern Europe: The metropolitan/non-metropolitan divide in the pre- and post-crisis era. *Eur. Plan. Stud.* **2018**, *26*, 256–278. [[CrossRef](#)]
29. Gorzelak, G.; Goh, C. (Eds.) *Financial Crisis in Central and Eastern Europe. From Similarity to Diversity*; Barbara Budrich Publishers: Leverkusen, Germany, 2010.
30. Psycharis, Y.; Kallioras, D.; Pantazis, P. *Regional Inequalities in Central and Eastern European Countries: The Role of Capital Regional and Metropolitan Areas, In Economic Development and Financial Markets Latest Research and Policy Insights from Central and Southeastern Europe*; Śliwiński, A., Karasavoglou, A., Polychronidou, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; pp. 3–20. [[CrossRef](#)]
31. Łuczka, T.; Przepióra, P. Regional Determinants of Efficiency Growth of Small and Medium-Sized Enterprises. Evidence from Poland. *J. Entrep. Manag. Innov.* **2012**, *8*, 138–154.
32. Wach, K. Regional Barriers and Stimuli of SMEs Development in Southern Poland. *SSRN Electron. J.* **2007**. [[CrossRef](#)]
33. Lang, T. Socio-economic and political responses to regional polarisation and socio-spatial peripheralisation in central and Eastern Europe: A research agenda. *Hung. Geogr. Bull.* **2015**, *64*, 171–185. [[CrossRef](#)]
34. Zbiorowa, P. *Raport o Stanie Sektora Małych i Średnich Przedsiębiorstw w Polsce*; PARP: Warsaw, Poland, 2019.
35. Kohonen, T. *Self-Organization and Associative Memory*; Springer: New York, NY, USA, 1988.