

# Localization Decisions of Entrepreneurs: The Role of Path Dependency and Market Forces

Korneliusz Pylak <sup>1</sup>, Dariusz Majerek <sup>1</sup>

<sup>1</sup> Lublin University of Technology, Nadbystrzycka 38d St., 20–618, Lublin, Poland

korneliusz.pylak@pollub.pl

**Abstract.** The purpose of this research is to determine the role of path dependency and market forces in the localization decisions of entrepreneurs from different industries. We hypothesize that most industries develop new entities based on the number of companies from the same industry that already exist in a region. We also hypothesize that entrepreneurs create new entities based on related industries operating within the same knowledge pool. To test these hypotheses, we used the machine learning decision tree method. The input variables are the number of companies from 86 industries located in 2,531 communities in Poland in 2009. The target values are the number of new companies from these industries created in the years 2009–2015. The principal results show that localization decisions are mostly based on demand and supply industries, in which manufacturing industries play crucial role. Path dependency appears in less than half of the industries' models and thus is not the main factor influencing decisions regarding the creation of new companies. The highest share of path-dependent industries is in manufacturing sector, but the degree of the dependence is lower than in the service sector. The service sector seems to be the least path-dependent, as services usually serve other industries. Competition in industries is a rare factor in new company creation; however, if it appears, it usually shrinks the industry.

## 1. Introduction

The purpose of this research is to determine the role of path dependency and market forces in the localization decisions of entrepreneurs from different industries in Poland. According to the business literature, numerous factors affect entrepreneurs deciding to start new companies or open new branches of existing ones, including regional or local labour markets, infrastructure, the environment, and economic and social stability. However, the literature rarely considers factors related to path dependency [1] and market forces [2], as it focuses on factors attracting big or foreign investors to new destinations [3], which usually change the local economy's structure and renew the local or regional development path. Large new investments are one of the most significant mechanisms allowing local or even regional economies to break out of path dependency. However, little is known about the localization decisions that relatively small and local, home-grown companies make, although they create the most jobs [4]. Companies make many of these decisions in every community, and these decisions are likely to be path-dependent and market-driven. In fact, path dependency intersects with market forces, as suppliers, competitors, and clients may be path-dependent to varying extents. Low barriers to market entry and the ease with which substitutes can be introduced can also lead to path dependency. Thus, these two phenomena are the subject of this study, and their roles in localization decisions will be analyzed in



detail. Because factors affecting company creation differ across industries [5], we analyze each industry separately.

The paper is structured as follows: The second section explores and clarifies the theoretical background of entrepreneurs' localization decisions and the roles of path dependency and market forces in the decision process. The third section lays out the empirical design of the research, including assumptions, hypotheses, data, and measurement. The fourth section presents and discusses the empirical results of the analysis regarding both path dependency and market forces. The fifth section offers a conclusion regarding factors influencing entrepreneurs' localization decisions.

## **2. Factors influencing entrepreneurs' localization decisions**

The factors influencing entrepreneurs' localization decisions are diverse and can be divided into several groups: 1) five market forces [6]; 2) resources' factors, including specific human, material, intellectual, and financial resources; 3) socioeconomic factors, including cognitive, behavioural, entrepreneurial, and social factors [2]; 4) institutional factors related to innovation systems, research, and business environmental institutions; and 5) global factors, including global industry cycles, global value chains, and international relations [see more in 7]. While the first group of factors can directly lead to the creation of a new company, the rest of the factors act as catalysts or retardants of the company creation process. Thus, we focus on the first group of factors.

### *2.1. Market forces*

Companies operate in an environment that Porter [6] described according to five competitive market forces. These market forces include the bargaining power of suppliers and customers, which may facilitate the creation of new companies within an industry or even create a new industry, if, for example, a company introducing new products needs a specific supplier that does not exist in the neighbourhood (clearly, the company may also import materials and semi-finished products from outside the neighbourhood). The other forces refer to competition within the industry and include the threats of new entrants and the introduction of substitute products or services. These forces can either hamper the development of companies because of high competition in a given industry or facilitate their development through lowering entry barriers. However, low entry barriers may increase rivalry within an industry and thus limit new company creation. Limited demand or supply potential may also increase rivalry in the industry.

The demand for new products or services is also a common trigger of new industry creation [6], but other factors are needed for such creation, and the creation process is very complex [7]. First, the cognitive needs should be fulfilled, as the customers demand new products or services and learn how and for what to use them. Then, a new industry emerges when entrepreneurs succeed in mobilizing human, material, financial, and intangible resources. However, this emergence depends on the skills of the entrepreneurs and workers and the state of the economy in general [2].

### *2.2. Path dependency*

Path dependency, introduced by David [8] and Arthur [9], means that an industry is locked on the development trajectory shaped by the industry's history. Although path dependency is a non-ergodic process that evolves unpredictably and irregularly, the renewal of an industry is unlikely unless some improbable event or shock disrupts the industry's locked-in pattern [10]. In most cases, companies and their business environments unknowingly strengthen the self-reinforcing mechanism that deepens industries' path dependency [11]. The self-reinforcing mechanism is created through natural resource endowment, technical relatedness, and thus the quasi-irreversibility of investments in technology [8, 12]. Subsequently, similarities in technologies and production create economies of scale and specialization on one hand and routine, and copying of others on the other [8, 11], and the industry becomes less creative and more resistant to changes. Therefore, the renewal of an existing industry or the creation of a new industry in a path-dependent environment is almost impossible. Thus, a regional economy can create new companies mostly in related industries based on similar knowledge pools [13].

In other words, path dependency encourages regional economies to continue specializing within related industries.

The creation of new industries is challenging for entrepreneurs because entrepreneurs rarely have the cognitive and socio-political legitimacy [2] to engage in such creation. Entrepreneurs willing to start a business in a new industry need to mobilize all tangible and intangible resources and convince customers to use their new products or services (unless customers are the driving force of the new industry). However, the resources are rarely available for new ideas since the uncertainty surrounding new ideas limits the resources available to bring them to fruition. In particular, workers will not absorb new knowledge unrelated to existing knowledge pools in the industry because of their lack of skills [14]. Also, while entrepreneurs' previous entrepreneurial, technical, or managerial experience clearly facilitates the recognition of new, innovative business opportunities, they usually create new companies in familiar industries [15]. New industries may emerge after a positive shock or event in the local environment, such as the creation of new transnational corporation's (TNS) branch or factory in the neighbourhood. The TNS can both belong to an existing industry in local economy and be the first representative of a new industry. Because a TNS often needs to cooperate with local companies, its suppliers may also belong to existing or new industries [3]. Cooperating with a TNS is a reliable and low-risk venture; thus entrepreneurs willingly risk entering completely new industrial areas [2].

In this study, we do not judge whether the localization decisions made within industries are beneficial for the local economy, as the debate on Marshallian and Jacobs's externalities is still inconclusive [16]. Path dependency can benefit regional and technological development as it facilitates continuous progress within the existing trajectory and thus higher predictability of outcomes, lower risk of failures, easier learning abilities, routinization of innovation processes, easier availability and production of components and complementary products, and, as a result, economies of scale [17]. These findings concerning predictability and efficiency are contrary to the position of Sydow et al. [11]. However, in the long term, cognitive and economic constraints caused by path dependency may cause new opportunities from beyond the path to be missed and thus such opportunities may not be beneficial [17].

### 3. Empirical design

#### 3.1. General approach

Our aim is to build models for industries to assess the influence of path dependency and market forces on the creation of new companies in Poland. The appearance of specific industries in the model and their influence on the change in the number of new companies can reflect both path dependency and market forces. A path-dependent industry is mostly likely to induce the emergence of a new company within the same industry. In our understanding, the more influential the existing companies of the industry are in the creation of new companies, the more path-dependent the industry is [13]. A path-dependent region tends to create new companies in existing and dominating industries and thus to increase the share of these industries in its economy and specialize [18]. Based on the current literature, we hypothesize that *(H.1) path dependency is the main factor in the creation of new companies in Polish industries.*

However, the appearance of an industry in the model does not necessarily mean the industry is path-dependent. The companies from the industry may cause a decrease in the number of new companies, indicating high rivalry in the industry. Although Klepper and Graddy [5] claim that in the second phase of industry development, nearly every industry experiences shakeout in the number of companies, followed by a stabilization phase. The size and scope of the shakeout usually depends on the type of industry, and the same industry across the country may be affected to varying degrees. If companies in the industry cause a significant decrease in the number of companies in the model, the industry may be exposed to fierce competition. Thus, *(H.2) industries with high internal competition experience shrinkage as the number of new companies decreases.*

The threat of new entrants to the industry is reflected by the scale of the number of new companies created in the model. The more companies are created, the easier entrance to the industry is. The number of new entrants is inseparable from the average size of the company, as the entrants depend on economies of scale, capital requirements, and product differentiation to overcome other product royalties

[6]. Other market forces related to demand, supply and substitutes are reflected by different industries appearing in a model for a given industry. We hypothesize that *(H.3) suppliers and clients are the most common predictors of new company creation in every industry, and the majority of these predictors belong to service industries*, as Pylak and Majerek [18] state.

### 3.2. Data and measurement

The aim of the analysis is to test the influence of path dependency and market forces on the creation of new companies in each Polish industry. One possible method for predicting these influences is a multiple regression; however, the structure of the predictors' relationships is complex, and the redundancy of endogenous variables means that most of them are collinear. Thus, as an alternate, we used decision trees as a machine learning method, as it does not require the structure of relationships to be known, and the redundancy of predictors is not an issue [19]. First, we computed the variable importance for each predictor in the models by assessing its overall improvement to the optimization criteria. Importance is a measure of the total predictor influence on the target variable. Then, we included each categorical predictor in the models as a single entity; thus each model determines how to group or split the values [20]. By decreasing the complexity parameter, the decision tree gains a lower predictive performance but higher generalization abilities, and thus only crucial predictors are chosen. However, our analysis covered not only all predictors that appeared with non-zero importance in the models but also predictors chosen by the models to be included in the trees because we noticed that the models sometimes rejected predictors with high importance and that fit well in the economic value chain.

To validate the predictions, we randomly divided the sample into two groups: a training group and a test group, which account for 75% and 25% of the records, respectively. The first group was used to train the model, and the second group was used to validate the model's prediction performance through analyzing the mean-square errors (*mse*) given by (1)

$$mse = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (1)$$

where  $N$  is the sample size,  $y_i$  is the observed value, and  $\hat{y}_i$  is the fitted value from the model.

To test the hypotheses, we used the number of companies from 86 industries located in 2,531 communities in Poland in 2009 as input variables and the number of new companies from these industries created in the years 2009–2015 as target values, creating 86 models (trees). Decision trees allow us to predict how many new companies from each industry will be created in communities given the precise number of companies from the same and related industries and the number of communities in which these relationships occur. Thus, the trees indicate which elements of economic structure influence the creation of new companies from every industry to the greatest extent.

## 4. Results and discussion

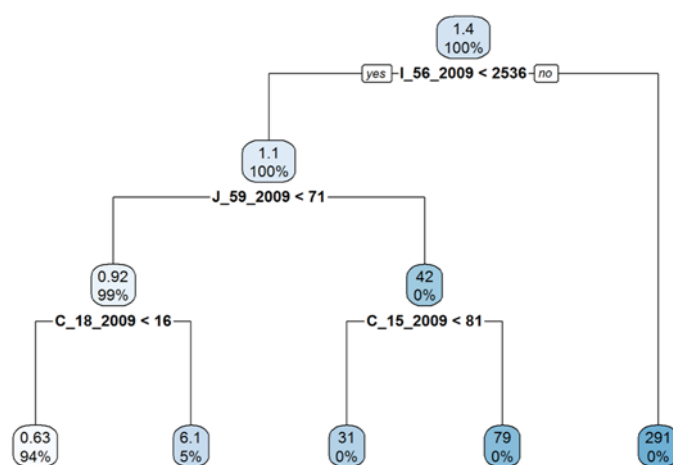
### 4.1. Path dependency and rivalry in the industry

First, we analyzed the appearance of industries in their models as predictors of the creation of new companies to determine the influence of path dependency and rivalry on the industry. Table 1 shows that nearly 60% of models contained the same industry as a potential predictor, particularly in manufacturing (75%) but to the least extent in services (48%). However, the rankings of the same industries are much higher in the service sector (7.0) than in manufacturing (11.1). Then, we analyzed decision trees in which the same industry was chosen as a predictor. The shares of such models were notably lower than in the case of potential predictors but still proportionate. Finally, we investigated the influence of existing companies from the analyzed industry on the creation of new companies in the same industry. The positive influence of path dependency was noted in 33% of all industries, including 46% of manufacturing industries, 44% of other industries, and 22% of service industries. Knowledge intensity does not play a crucial role in the process, as there is no correlation between the share of path-dependent industries and knowledge intensity.

**Table 1.** Path dependency and market forces in entrepreneurs' localization decision trees.

Type of industry	[n]	The share of industries with the same industry as a potential predictor (with significant importance)	Average rank of the same industry as a predictor (1 = the most important and 86 = the least important)	The share of industries in which the same industry is chosen as a predictor in the decision tree	The share of path-dependent industries	The share of industries with internal rivalry decreasing the number of new companies
All industries	86	59.3%	8.2	41.9%	32.6%	12.8%
Manufacturing, including:	24	75.0%	11.1	58.3%	45.8%	12.5%
low-technology industries	11	81.8%	9.8	63.6%	45.5%	18.2%
medium-low-technology industries	6	83.3%	13.8	66.7%	66.7%	0.0%
medium-high-technology industries	5	60.0%	13.3	60.0%	40.0%	20.0%
high-technology industries	2	50.0%	3.0	0.0%	0.0%	0.0%
Services, including:	46	47.8%	7.0	30.4%	21.7%	13.0%
high-tech knowledge-intensive services (KIS)	6	50.0%	13.0	33.3%	33.3%	0.0%
market KIS	9	11.1%	6.0	11.1%	11.1%	0.0%
financial KIS	3	100.0%	3.3	66.7%	33.3%	66.7%
other KIS	11	45.5%	9.2	18.2%	18.2%	9.1%
less-knowledge-intensive services	17	58.8%	5.4	41.2%	23.5%	17.6%
Other industries	16	68.8%	5.6	50.0%	43.8%	12.5%

Competition in the industries is an important factor for nearly 13% of industries, affecting all sectors equally. In knowledge-intensive high-tech and market services, rivalry does not play a role, but in financial KIS, 2 of 3 industries are shaped by severe competition (K\_65: insurance, reinsurance and pension funding, except compulsory social security, and K\_66: activities auxiliary to financial services and insurance activities). Interestingly, in most cases (7 out of 11), rivalry in the industry shrank the industry, meaning that the number of new companies tended to decrease over the period (see Appendix 1).

**Figure 1.** The decision tree for the C\_10 industry (manufacture of food products)

#### 4.2. Supplier and client industries and new entrants

We also analyzed the industries serving as potential suppliers and clients for a given industry. Table 2 shows the shares of industries in which new companies are co-created by supplier and client industries.

We could easily find at least one supplier or client industry with non-zero importance as a potential predictor in every model, although not every natural supplier or client was found to be a potential predictor. Less numbers of suppliers and clients were chosen in trees, however in some cases, they played the most important role in company creation process (see, for example, figure 1). The creation of new companies in the manufacture of food products is influenced the most by the client industry (I\_56: food and beverage service activities). In general, the creation of companies in all manufacturing industries is influenced by at least one supplier and one client industry. In the service sector, the share of industries influenced by supplier and client industries is lower, but these industries are less dependent on institutional suppliers and clients and more focused on people than manufacturing industries.

**Table 2.** Influence of threats of new entrants and the power of suppliers and clients on entrepreneurs' localization decisions.

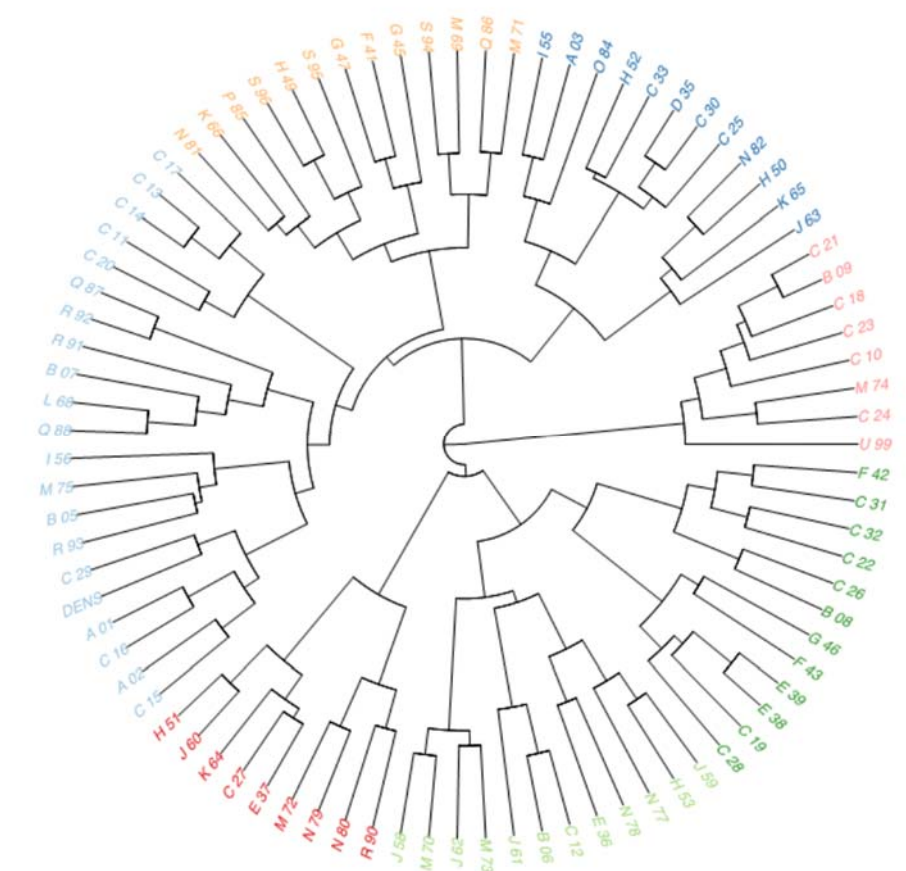
Type of industry	[n]	The most likely average number of companies created in the model	The share of industries in which new companies are co-created by supplier industries (at least one industry with potential importance)	The share of industries in which new companies are co-created by client industries (at least one industry with potential importance)
All industries	86	1.150	90.7%	83.7%
Manufacturing, including:	24	0.282	100.0%	95.8%
low-technology industries	11	0.230	100.0%	90.9%
medium-low-technology industries	6	0.574	100.0%	100.0%
medium-high-technology industries	5	0.118	100.0%	100.0%
high-technology industries	2	0.097	100.0%	100.0%
Services, including:	46	1.672	84.8%	73.9%
high-tech KIS	6	1.102	83.3%	50.0%
market KIS	9	1.814	55.6%	100.0%
financial KIS	3	0.097	100.0%	100.0%
other KIS	11	1.603	100.0%	45.5%
Less-knowledge-intensive services	17	2.121	88.2%	82.4%
Other industries	16	0.950	93.8%	93.8%

Threats of new entrants are most common in the service sector, especially in less-knowledge-intensive services and market services. The process of new companies entering the industry must be industry-specific, as the process does not depend on knowledge intensity. In the manufacturing sector, it is easier to create more companies in medium-low-technology industries than in low-technology industries, though high-technology industries most commonly create only 0.097 companies over the period.

#### 4.3. The most influential industries in company creation

Finally, to test if the service sector plays a crucial role in company creation, we analyzed the importance values of predictors in all the models, including predictors' number of appearances in models and values compared to other predictors (reflected by their ranks in the models). Then, we conducted cluster analysis to see if the most influential industries belonged to the service sector. The findings are shown in figure 2. The most influential cluster (in bright red) consists of B\_09, C\_10, C\_18, C\_21, C\_23, C\_24, M\_74, and U\_99, indicating that there are five manufacturing industries and only two service industries.





**Figure 2.** Cluster analysis of most influential predictors in the models

## 5. Conclusions

The aim of the analysis was to test the influence of path dependency and market forces on the creation of new companies in a variety of industries. Three hypotheses were tested using the decision tree method, with new findings explained in detail. The principal results show the diverse role of path dependency in the process of new company creation. The highest share of path-dependent industries is in the manufacturing sector, but the dependence of individual industries is weaker than in the service sector. The share of path-dependent industries in the service sector is the lowest, as services usually serve many other industries and thus service sector is more flexible and adaptive than other sectors. However, in general, no more than half of the industries are path-dependent, which means that path dependency does not play the main role in company creation (hypothesis H.1 is not confirmed).

Competition in the industry shapes new company creation in 12% of industries, and most of these industries are shrinking due to high rivalry, confirming hypothesis H.2. In turn, new company creation is shaped mostly by demand and supply. Cluster analysis revealed that the most influential industries belong to the manufacturing sector; thus hypothesis H.3 is not confirmed. Also, the spread of the potential predictors in the models (the number of predictors) shows that the manufacturing sector has the highest related variety of industries (28.5 industries per model average), while the service sector has only 19.7 industries per model average and other industries have 24.2.

## References

- [1] R. Boschma, A. Minondo, and M. Navarro, “Related variety and regional growth in Spain,” *Papers in Regional Science*, vol. 91, no. 2, pp. 241–256, 2012.

- [2] H. E. Aldrich and C. M. Fiol, "Fools rush in? The institutional context of industry creation," *Academy of Management Review*, vol. 19, no. 4, pp. 645–670, 1994.
- [3] D. Castellani and A. Zanfei, "Multinational experience and the creation of linkages with local firms: Evidence from the electronics industry," *Cambridge Journal of Economics*, vol. 26, no. 1, pp. 1–25, 2002.
- [4] D. Neumark, B. Wall, and J. F. Zhang, "Do small businesses create more jobs? New evidence for the United States from the national establishment time series," *Review of Economics and Statistics*, vol. 93, no. 1, pp. 16–29, 2011.
- [5] S. Klepper and E. Graddy, "The evolution of new industries and the determinants of market-structure," *Rand Journal of Economics*, vol. 21, no. 1, pp. 27–44, 1990.
- [6] M. E. Porter, "How competitive forces shape strategy," *Harvard Business Review*, vol. 57, no. 2, pp. 137–145, 1979.
- [7] J. H. Chen and Y. J. Chen, "New industry creation in less developed countries—the case of the Taiwanese flat panel display industry," *Innovation-Management Policy & Practice*, vol. 17, no. 2, pp. 250–265, 2015.
- [8] P. A. David, "Clio and the economics of QWERTY," *American Economic Review*, vol. 75, no. 2, pp. 332–337, 1985.
- [9] W. B. Arthur, "Competing technologies, increasing returns, and lock-in by historical events," *Economic Journal*, vol. 99, no. 394, pp. 116–131, 1989.
- [10] M. Henning, E. Stam, and R. Wenting, "Path dependence research in regional economic development: Cacophony or knowledge accumulation?" *Regional Studies*, vol. 47, no. 8, pp. 1348–1362, 2013.
- [11] J. Sydow, G. Schreyoegg, and J. Koch, "Organizational path dependence: Opening the black box," *Academy of Management Review*, vol. 34, no. 4, pp. 689–709, 2009.
- [12] R. Martin, "Regional economies as path-dependent systems: Some issues and implications," in *Handbook of Regional Innovation and Growth*, P. Cooke, B. Asheim, R. Boschma, R. Martin, D. Schwartz, and F. Todtling, Eds. Cheltenham: Edward Elgar, pp. 198–210, 2011.
- [13] R. Boschma, A. Minondo, and M. Navarro, "The emergence of new industries at the regional level in Spain: A proximity approach based on product relatedness," *Economic Geography*, vol. 89, no. 1, pp. 29–51, 2013.
- [14] R. Boschma and S. Iammarino, "Related variety, trade linkages, and regional growth in Italy," *Economic Geography*, vol. 85, no. 3, pp. 289–311, 2009.
- [15] A. Alvarez-Herranz, P. Valencia-De-Lara, and M. P. Martinez-Ruiz, "How entrepreneurial characteristics influence company creation: A cross-national study of 22 countries tested with panel data methodology," *Journal of Business Economics and Management*, vol. 12, no. 3, pp. 529–545, 2011.
- [16] K. Frenken, F. Van Oort, and T. Verburg, "Related variety, unrelated variety and regional economic growth," *Regional Studies*, vol. 41, no. 5, pp. 685–697, 2007.
- [17] M. Guerzoni, T. T. Aldridge, D. B. Audretsch, and S. Desai, "A new industry creation and originality: Insight from the funding sources of university patents," *Research Policy*, vol. 43, no. 10, pp. 1697–1706, 2014.
- [18] K. Pylak and D. Majerek, "Impact of the service sector on the creation of companies in Poland," *Procedia Economics and Finance*, vol. 24, pp. 523–532, 2015.
- [19] M. Kuhn, "Building predictive models in R using the caret package," *Journal of Statistical Software*, vol. 28, no. 5, pp. 1–26, 2008.
- [20] M. Kuhn and K. Johnson, *Applied predictive modeling*. New York: Springer, 2013.



## Appendix

**Table 3.** Path dependency and market forces in entrepreneurs' localization decision trees.

Industry	Sector <sup>a</sup>	Knowledge intensity <sup>b</sup>	The same sector as a potential predictor (with significant importance)	Rank of the predictor	The predictor is chosen in the decision tree	The industry is path-dependent	The rivalry inside the industry is decreasing the number of new companies	The most common number of companies created	The share of industries co-created by suppliers (at least one industry with pot. importance)	The share of industries co-created by clients (at least one industry with potential importance)
A_01: Crop and animal production, hunting and related service activities	O	O	+	1	+	–	+	–1.2	+	+
A_02: Forestry and logging	O	O	+	1	+	+	–	0.43	+	+
A_03: Fishing and aquaculture	O	O	+	1	+	+	+	0.071	+	+
B_05: Mining of coal and lignite	O	O	+	1	+	+	–	0.000551	+	+
B_06: Extraction of crude petroleum and natural gas	O	O	–	–	–	–	–	0.0075	+	+
B_07: Mining of metal ores	O	O	+	8	+	+	–	0.0027	+	+
B_08: Other mining and quarrying	O	O	+	20	+	+	–	0.37	+	+
B_09: Mining support service activities	O	O	+	2	–	–	–	0.079	+	+
C_10: Manufacture of food products	M	LT	+	4	–	–	–	0.63	+	+
C_11: Manufacture of beverages	M	LT	–	–	–	–	–	0.022	+	+
C_12: Manufacture of tobacco products	M	LT	+	29	+	+	–	0.0068	+	–
C_13: Manufacture of textiles	M	LT	+	4	+	+	–	0.41	+	+
C_14: Manufacture of wearing apparel	M	LT	+	1	+	–	+	–0.19	+	+
C_15: Manufacture of leather and related products	M	LT	+	1	+	–	+	–0.064	+	+
C_16: Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	M	LT	+	8	+	+	–	0.34	+	+
C_17: Manufacture of paper and paper products	M	LT	–	–	–	–	–	0.28	+	+
C_18: Printing and reproduction of recorded media	M	LT	+	28	–	–	–	0.3	+	+
C_19: Manufacture of coke and refined petroleum products	M	MLT	+	7	+	+	–	0.026	+	+
C_20: Manufacture of chemicals and chemical products	M	MHT	–	–	–	–	–	0.17	+	+
C_21: Manufacture of basic pharmaceutical products and pharmaceutical preparations	M	HT	+	3	–	–	–	0.024	+	+
C_22: Manufacture of rubber and plastic products	M	MLT	+	18	+	+	–	0.31	+	+
C_23: Manufacture of other non-metallic mineral products	M	MLT	+	9	+	+	–	0.11	+	+
C_24: Manufacture of basic metals	M	MLT	–	–	–	–	–	0.1	+	+
C_25: Manufacture of fabricated metal products, except machinery and equipment	M	MLT	+	34	–	–	–	1.3	+	+
C_26: Manufacture of computer, electronic and optical products	M	HT	–	–	–	–	–	0.17	+	+
C_27: Manufacture of electrical equipment	M	MHT	–	–	–	–	–	0.091	+	+
C_28: Manufacture of machinery and equipment n.e.c.	M	MHT	+	22	+	–	+	0.23	+	+
C_29: Manufacture of motor vehicles, trailers and semi-trailers	M	MHT	+	2	+	+	–	0.052	+	+
C_30: Manufacture of other transport equipment	M	MHT	+	16	+	+	–	0.047	+	+
C_31: Manufacture of furniture	M	LT	+	7	+	+	–	0.57	+	+
C_32: Other manufacturing	M	LT	+	6	+	+	–	0.23	+	+
C_33: Repair and installation of machinery and equipment	M	MLT	+	1	+	+	–	1.6	+	+

Industry	Sector <sup>a</sup>	Knowledge intensity <sup>b</sup>	The same sector as a potential predictor (with significant importance)	Rank of the predictor	The predictor is chosen in the decision tree	The industry is path-dependent	The rivalry inside the industry is decreasing the number of new companies	The most common number of companies created	The share of industries co-created by suppliers (at least one industry with pote. importance)	The share of industries co-created by clients (at least one industry with potential importance)
D_35: Electricity, gas, steam and air conditioning supply	O	O	-		-	-	-			
E_36: Water collection, treatment and supply	O	O	-		-	-	-	0.055	+	+
E_37: Sewerage	O	O	+	13	+	+	-	0.26	+	+
E_38: Waste collection, treatment and disposal activities; materials recovery	O	O	-		-	-	-	0.45	+	+
E_39: Remediation activities and other waste management services	O	O	-		-	-	-	0.073	-	-
F_41: Construction of buildings	O	O	+	12	-	-	-	2.1	+	+
F_42: Civil engineering	O	O	+	2	+	+	-	1.5	+	+
F_43: Specialised construction activities	O	O	+	1	-	-	-	10	+	+
G_45: Wholesale and retail trade and repair of motor vehicles and motorcycles	S	LKIS	-		-	-	-	5	-	+
G_46: Wholesale trade, except of motor vehicles and motorcycles	S	LKIS	+	2	-	-	-	4.9	+	+
G_47: Retail trade, except of motor vehicles and motorcycles	S	LKIS	+	1	+	-	+	-4	+	+
H_49: Land transport and transport via pipelines	S	LKIS	+	7	+	-	+	-0.49	+	+
H_50: Water transport	S	KISM	+	6	+	+	-	0.058	+	+
H_51: Air transport	S	KISM	-		-	-	-	0.044	+	+
H_52: Warehousing and support activities for transportation	S	LKIS	+	7	-	-	-	0.85	+	+
H_53: Postal and courier activities	S	LKIS	+	8	-	-	-	0.41	+	+
I_55: Accommodation	S	LKIS	+	7	+	+	-	1.2	+	+
I_56: Food and beverage service activities	S	LKIS	+	7	+	+	-	0.87	+	+
J_58: Publishing activities	S	KISO	-		-	-	-	0.18	+	+
J_59: Motion picture, video and television programme production, sound recording and music publishing activities	S	KISH	-		-	-	-	0.52	+	-
J_60: Programming and broadcasting activities	S	KISH	+	15	+	+	-	0.032	+	-
J_61: Telecommunications	S	KISH	-		-	-	-	0.35	+	+
J_62: Computer programming, consultancy and related activities	S	KISH	+	7	+	+	-	4.5	+	-
J_63: Information service activities	S	KISH	+	17	-	-	-	1	+	+
K_64: Financial service activities, except insurance and pension funding	S	KISF	+	1	-	-	-	0.9	+	+
K_65: Insurance, reinsurance and pension funding, except compulsory social security	S	KISF	+	8	+	+	+	0.19	+	+
K_66: Activities auxiliary to financial services and insurance activities	S	KISF	+	1	+	-	+	-0.8	+	+
L_68: Real estate activities	S	LKIS	-		-	-	-	8.1	+	+
M_69: Legal and accounting activities	S	KISM	-		-	-	-	6.6	+	+
M_70: Activities of head offices; management consultancy activities	S	KISM	-		-	-	-	3.7	-	+
M_71: Architectural and engineering activities; technical testing and analysis	S	KISM	-		-	-	-	2.1	+	+
M_72: Scientific research and development	S	KISH	-		-	-	-	0.21	-	+
M_73: Advertising and market research	S	KISM	-		-	-	-	2.5	+	+
M_74: Other professional, scientific and technical activities	S	KISM	-		-	-	-	0.71	-	+

Industry	Sector <sup>a</sup>	Knowledge intensity <sup>b</sup>	The same sector as a potential predictor (with significant importance)	Rank of the predictor	The predictor is chosen in the decision tree	The industry is path-dependent	The rivalry inside the industry is decreasing the number of new companies	The most common number of companies created	The share of industries co-created by suppliers (at least one industry with pote. importance)	The share of industries co-created by clients (at least one industry with potential importance)
M_75: Veterinary activities	S	KISO	+	–	+	+	–	0.59	+	+
N_77: Rental and leasing activities	S	LKIS	–	–	–	–	–	1.6	+	+
N_78: Employment activities	S	KISM	–	–	–	–	–	0.44	–	+
N_79: Travel agency, tour operator and other reservation service and related activities	S	LKIS	–	–	–	–	–	0.34	+	–
N_80: Security and investigation activities	S	KISM	–	–	–	–	–	0.17	–	+
N_81: Services to buildings and landscape activities	S	LKIS	–	–	–	–	–	2.5	+	+
N_82: Office administrative, office support and other business support activities	S	LKIS	–	–	–	–	–	2.2	+	+
O_84: Public administration and defence; compulsory social security	S	KISO	+	6	–	–	–	0.082	+	–
P_85: Education	S	KISO	+	18	–	–	–	6.3	+	–
Q_86: Human health activities	S	KISO	–	–	–	–	–	7.9	+	+
Q_87: Residential care activities	S	KISO	–	–	–	–	–	0.18	+	+
Q_88: Social work activities without accommodation	S	KISO	–	–	–	–	–	0.42	+	–
R_90: Creative, arts and entertainment activities	S	KISO	–	–	–	–	–	0.54	+	–
R_91: Libraries, archives, museums and other cultural activities	S	KISO	+	14	+	+	–	0.088	+	+
R_92: Gambling and betting activities	S	KISO	+	7	–	–	+	0.056	+	–
R_93: Sports activities and amusement and recreation activities	S	KISO	–	–	–	–	–	1.3	+	–
S_94: Activities of membership organisations	S	LKIS	+	13	+	+	–	9.4	+	+
S_95: Repair of computers and personal and household goods	S	LKIS	+	1	+	–	+	–0.026	+	+
S_96: Other personal service activities; T_97: Activities of households as employers of domestic personnel; T_98: Undifferentiated goods- and services-producing activities of private households for own use	S	LKIS	–	–	–	–	–	3.2	+	–
U_99: Activities of extraterritorial organisations and bodies	S	LKIS	+	1	+	+	–	0.0054	–	–

<sup>a</sup> M – manufacturing, S – services, O – other activities, <sup>b</sup> LT – low technology, MLT – medium-low technology, MHT – medium-high technology, HT – high technology, KISH – knowledge-intensive high-tech services, KISM – knowledge-intensive market services, KISF – knowledge-intensive financial services, KISO – other knowledge-intensive services, LKIS – less-knowledge-intensive services