Localization Economies in the Bulgarian Production Sector: Empirical Evidence from the Pharmaceutical and Oil Production Industries

*** Final Thesis Paper ***

MSc Final Thesis

MSc Economics of Management and Organization

Erasmus School of Economics

Supervisor: Prof. Sasha Kapoor

Second Assessor: Prof. Laura Hering

Prepared by: Yulian Silyanovski (455593ys)

11th of September 2018

ERASMUS UNIVERSITEIT ROTTERDAM

Localization Economies in the Bulgarian Production Sector: Empirical Evidence from the Pharmaceutical and Oil Production Industries

Abstract: The purpose of this paper is to study the effect of localization economies on firms' performance in an attempt to answer the following research question:

RQ: *To what extent is there localization economies in the pharmaceutical (and oil) production industry in Bulgaria and how this affects the within-industry firm's (financial) performance in the period 2010-2016?*

In the way of answering the above-imposed question, this paper examines the effectiveness of within-industry clustering on firms' financial performance by focusing on the pharma and oil production industries in Bulgaria. Importantly, the dataset contains 7 years of key financial performance indicators and a localization employment quotient variable which serves as a proxy of within-industry clustering. Leaning on these variables, it is possible to investigate to what extent is there a localization phenomenon in the production industries in Bulgaria and how this grouping affects the financial performance of the firms.

The findings of this research suggest a somewhat dubious effect of a within-industry constellation for the tested production companies. Initially, considering the oil production data sample, no significant effect from industry grouping on firm's performance is to be observed (i.e. both sales and profit indicators are explained by other factors). On the other hand, once the dataset is filtered to pharmaceutical companies, a positive and significant association of firm's localization with profit figures is to be distinguished. Yet, referring to the same industry, no statistically significant relation between localization and sales performance is to be identified. Similarly, no particular difference between localized and less localized (i.e. more isolated) industry-leavers and/or entrants is to be revealed either. These discoveries remain robust throughout all additional panel data investigations.

Overall, the research encloses that for some production industries in Bulgaria, a greater level of within industry grouping is positively related to annual profit records, however, such grouping does not necessarily lead to healthier sales figures. A solid explanation for this is the fact that clustered pharma production companies experience a more efficient utilization of inputs - a process which is argued to partially stem from the positive externalities, gained by the companies in the more localized geographical regions.

Notably, due to insufficient sources of information, similar cluster investigations are absent for the Bulgarian market. Although this paper sets the ground for the need of similar research, a more in-depth analysis of the cluster economies in Bulgaria is required. Similar testing should be carried out for other industries in order to expropriate the true benefits of agglomeration economies in this geographical region.

Table of Contents

1. Introduction and Research Problem

Lone organizations, Localization of companies, Cluster of industries, Globalization of economies, Creation of ecosystems…[1]

The above-imposed naturally staggering list of keywords has become a widely discussed topic around the Globe and has attracted the attention (and actions) of many researchers, entrepreneurs, politicians, and CEOs of large corporations. In these lines, several economists, but before all M. Porter (1998), have identified the niche research area of clustered economies. Indeed, they all have developed concepts and studied theories which delineate the key influences of cluster economies upon the development of novel industries, the economy, and the society (Enright, 1996; Harrison, 1994; Harrison, Kelly, & Grant, 1996; Krugman, 1991; Rosenfeld, 1997; Storper, 1989).

Essentially, the theory behind this concept explains that clustered (or localized) firms tend to outperform organizations that are isolated (geographically). The key reason in support of this theory is that firms are to generate positive externalities, based on their geographic proximity, or hereafter referred to as localization of industries. (Kukalis, 2010; Porter, 1998). In other words, a company which operates in a relatively constellated geographical location, manages to generate superior profits (i.e. better performance) when compared to essentially the same but isolated organization. Researchers point out numerous of reasons why this might be the case. Some examples in support of this argument include superior infrastructure and means of communication, a higher density of customers or presence of clients' target groups (Kukalis, 2010; Porter, 1998). Moreover, such benefits have been proven to lead to an improved collaboration, increased R&D, lower spending and production costs, hence greater profits in the long run (Henderson, 1986; Ciccone & Hall, 1996; Saxenian, 1994; Bell, 2005, Jaffe, Trajtenberg and Henderson, 1993). Naturally, there are also contradictory findings which advocate that agglomeration of companies could come to a cost either. Some examples include knowledge spillovers, greater market competition (in the case of homogeneity of the firms), input firms and others (Kukalis, 2010).

Based on the above concept, the approach in this paper is twofold. By leaning on a self-collected pharma and oil production data from registered firms in Bulgaria, first, the author verifies which geographical areas correspond to what level of industry grouping. In the way of doing so,

[1] e.g. Silicon Valley in California.

the author calculates an employment localization quotient (LQ) which measures the overall employment factor of the tested industry, segregated by geographical regions.[2] Then, the estimated LQ variable is included in the main regression model to test the effect of localization on the performance of the firm. In other words, the accumulated effects of the within-industry clustering (i.e. the attached externalities which stem from a set of unobservable geographical factors) can be seen in the accounts of the firm. Thus, the key focus of the paper is to navigate the reader to the potential associated effects of this clustering, rather than explaining the exact reasons why companies cluster together.

Yet, the author of this paper focuses the research on the production companies only (with pharma and oil manufacturing being the tested ones). The key reasoning for this decision comes down to the fact that manufacturing and high-tech industries are reported to show more profound effects from specific localization (or cluster) economies (Kukalis, 2010; Cannina, Enz and Harrison, 2005, Porter, 1998). Moreover, with generating roughly $1/4^{th}$ of the total GDP, the Bulgarian market is largely driven by the B and E sectors (in which manufacturing and production industries are included).[3] Needless to say, according to NSI (2018) and Kapital (2018), with the largest (in revenue) company in Bulgaria (being in the oil industry) and the highest market cap valued company (being in the pharma industry), together these industries are (proportionately) part of the largest GDP generators on the Bulgarian market, which in turn, makes them a reasonable pick for this market investigation.

Finding an answer to the research question would enlighten the localization tendency of the production firms in the Bulgarian market. Considering the relevance of this research topic, the reported results would benefit companies in the investigated industries, as well as, will guide entrepreneurs whether to invest in such localized areas. Additionally, it might provide a novel way of looking into localized economies and can yield new directions of improved market investigation. Therefore, this paper will also be of captivating for any industry and university scholars, the Bulgarian government and businesses, as well as the international audience.

[2] A more precise depiction of the employment location quotient and its measurement can be found in Chapter 3 of this paper.

[3] The sectors information is extracted from the NACE Rev2 A10, in which manufacturing and oil production are reported to be included (NSI, 2018; Eurostat, 2018). A detailed GDP report for the period 2010 – 2017 can be found in Appendix A.

The structure of this research paper continues as follows: First, the author starts with a brief overview of the existing related literature to present the key hypothesis, followed by a summary of the data sample and an outline of the main regression estimation. Next, the core of this paper shall be delineated by an in-depth exploration of the empirical results and interpretation of the key findings. The paper will close in its discussion and conclusion section by referring to the main caveats and will provide relevant recommendations for future investigation.

2. Background Literature and Hypothesis Development

Digging deeper in literature, there are contradicting claims that drive firm's decisions to stagnate (i.e. cluster around other firms or industries) or to stay isolated. Surely, such strategic choices need to be somehow reflected with the short and long-term performance of the organization. Before looking at the specific case of whether and how localization economies relate to the (financial) performance of the firm, it is crucial to understand the underlying mechanisms.

Theoretically, scholars have identified four different concepts around the formation of cluster industries. These include: (1) the economic conditions of the geolocation of the cluster, (2) the rivalry within the cluster, (3) the formation and the progress of the cluster and (4) the impact of the cluster on the surrounding environment (Kukalis, 2010). Notably, the driving force behind these areas of investigation is the way in which the between-firms-interaction (and stagnation) help (or deteriorate) the individual firm's performance and how this affects the total (and individual) productivity, the surrounding environment, and the overall economic development of a geographic area. Nonetheless, considering the large scope of a full cluster investigation, in this research paper, the author has only focused on the observable results of firms' grouping (concepts (1) and (2)), rather than the reasons or consequences behind it (concepts (3) and (4)). Hence, a logical question that arises is whether an individual firm which operates close to others (in geographic terms), tends to benefit (financially) from doing so. In support of testing this argument, Kukalis (2010) explains that the financial performance of firms, which experience positive externalities from being in a specific cluster, should be visible from the accounts of the firm. Stemming from the logic above, the key interest of this paper constitutes in explaining the effectiveness of the firm's localization on her market performance.

Now, it is important to note that the concept of agglomeration (or cluster) economies has two different dimensions, i.e. localization and urbanization. The first (localization) refers to the effects of external economies of scale which are dependent on the performance and development of a single industry in a region (Marshall, 1980). The second (urbanization) refers to the externalities affected by the economic diversity (i.e. all industries and related parties) in each region or area (Jacobs, 1969). The aim of this paper is to test only the localization economies of the

pharmaceutical (and oil) production industry and its effect of on the financial performance of the firms.

Additionally, economists distinguish between internal (capabilities of a single firm) and external economies of scale (total performance of the tested industry within a single area). It is argued that externalities concern the benefits that firms in the same industry receive from being clustered together in one area (Marshall, 1890; Porter, 1998). Further, if an industry receives localization economies, then economic actors from the same industry are likely to gather together in a particular region which is specialized in a particular production of goods/services or any other closely interconnected activities (Marshall, 1890). This leads to a full exploitation of the external scale economies (Henderson, 2003). Examples of localization economies include: The Silicon Valley characterized by many but small software firms; Restaurants clustering near each other to gain from other outlets' customer outflow (Austin et al., 2005).

In order to examine the above-discussed relation, it is now important to provide the theoretical reasoning which wraps up the tested hypotheses. Yet, it was not before 1920, when A. Marshall proved to be the first scholar to effectively summarize the concept of "external economies of scale". By doing so, he managed to successfully stress on the important attributes of external factors on individual firms (Marshall, 1920). Fundamentally, Marshall argues that organizations develop a common (eco)system from which the resided firms benefit or detriment. Such effects are prone to stem from common resources such as infrastructure, factor conditions, government support, specialized labor force, and consumer demand (Porter, 1998). Nevertheless, it was not prior Michal Porter (1998), who managed to provide a proper depiction of cluster economies, how they are formed, and the vital importance they play on the overall economic development. In parallel to his study, some more concrete examples of constellation benefits are illustrated by numerous theoretical research. Evidences from the correlation between external effects and the firm's performance include: the importance of labor mobility and productivity (Almeida and Kogut, 1999), between-firms' interaction (Burt, 1992; Coleman, 1988; Saxenian, 1994), competitive advantage (Tallman et at., 2004), and the effect of spatial proximity or the dynamics of increasing returns (Krugman, 1991; Storper, 1989).

Summing the above literature, the concept of firm's stagnation occurs due to different reasons and leads towards positive (or negative) external and internal (dis)economies of scale. Now that

the theory is there, scholars tend to stand on polar viewpoints when it comes to the actual effects on the company's performance. By presenting the key benefits or drawbacks of localization economies, the following section portrays the tested hypothesis in this research.

On the one side, the majority of the existing literature expounds the concept that companies which operate in clustered areas tend to outperform isolated firms (Porter, 1998). The fact of the matter is that sitting in a cluster could actually lead to an increased firm's strategic competitiveness (hence performance), essentially caused by actors such as the presence of a constant peer-pressure (Porter, 1998). Porter also suggests that the overall effect is actually self-driven by firms and is led by "a healthy rivalry". This, in turn, generates increased total group (and individual) productivity, thus stronger pace of innovation and superior performance (Kukalis, 2010; Porter, 1998). Importantly, other reasons include: availability of resources, easier communication, coordination and networking (Kuklis, 2010). Furthermore, other factors, which blended together, yield positive externalities of sitting in a cluster, include: well-maintained logistics, presence of government assistance, greater overall demand (affected by the population growth or the presence of natural endowments (e.g. high-speed internet, water supply, etc.)) (Porter, 1998). It is also worth noting that firm's clustering can also lead to an increased labour productivity (Ciccone & Hall, 1996; Henderson, 1986) and/or to a reduction of input costs (Jaffe et al, 1993).

On the contrary, there are various research findings that companies do not always yield positive economies from sitting in a cluster. In fact, the effects could sometimes be neutral or in some cases, even negative (Kukalis, 2010). This is to say that in reality, a real difference between isolated and clustered companies could be absent. In support of this argumentation, Maskell (2001) implies that certain firms engage in destructive actions, which could sometimes deteriorate the long-term trust and cooperation of the whole industry. In order to avoid such non-cooperative misbehavior, some companies decide to operate in isolation (Krugman, 1991; Sexanian, 1994, Storper, 1989). Furthermore, proven by the empirical study of Shaver and Flyer (2000), it is found that firms who have a competitive advantage in technology, specialized human capital, and supplies and distribution, are more prone to locate away from others. Moreover, Holmes (1996) briefs the inclination of the cotton textile industry to isolate, whereas Stuart and Sorensen (2003) suggest that entrants in the biotechnology industry actually perform worse in a cluster. Nevertheless, such findings have been questioned due to insufficient evidence and are said to be informative only up

to the specificities of the tested industries and the examined geographical locations (Martin & Sunley, 2003; Kukalis, 2010).

Indeed, as proposed above, there are essential differences among the findings of the scholars. Theoretically speaking, however, the majority of the related literature suggests that the positive effects of clustering outnumber (both in quality and quantity) those against. Yet, it is also important to stress on the fact that grouping or clustering of firms occurs either because of (a) strategic business grouping (i.e. firms engage in an active and close cooperation (Granovetter, 1985)) or is formed by (b) industry clusters (i.e. firms are composed of autonomous decisions based on self-driven arm's length transactions (Porter, 1990, 1998, 2000)). Overall, the above-presented viewpoints suggest that regardless of the exact reasoning, cluster firms tend to get positive externalities stemming from the geographic proximity (propinquity). Now, the best way to evaluate the true positive effect of clustering on the performance of a firm is to look at the financial records of the company (e.g. profit and sales figures). These elements are proven to be instructive for the performance of the firm, hence should contain significant information about the effect that the localization plays on the tested industry (Kukalis, 2010).

In other words, leaning on the above argumentation that the agglomeration of firms leads to positive externalities such as access to recourses (common knowledge, R&D and innovation), improved networks and infrastructure, and ability to mimic the firm's behavior from successful examples, shall lead to an improved (financial) productivity of the firms that reside in such areas. This could be accompanied by lower operational costs and/or a higher labor productivity, which should reflect greater sales and essentially profits). The logical reasoning behind this concept leads to the development of the following hypothesis:

H1: The financial performance of production companies, which reside in a relatively more clustered region of within-industry localization, is comparatively better than those of isolated (less localized) firms in the same industry.

➔ Hence, localized (within-industry clustered) production firms (financially) outperform less-localized firms.

Furthermore, stemming from the above argumentation, it could be reasoned that the true effect of clustering can be caught when referring to the comparison between incubators in clustered versus isolated areas. Researchers have also found that diseconomies of scale can occur in the later stages of firm's development (life cycle), due to e.g. reduced skilled labor, increased barriers of

maintaining business, or a new government regulation (Audretsch & Feldman, 1996; Malecki, 1985). Therefore, due to the built-up common ecosystem, it could be reasoned that the effect of such diseconomies would be less profound in the clustered areas. Furthermore, due to expected spillover effects (of knowledge for instance), firms could learn how to adapt quicker to such uncontrollable changes and can perform better than the isolated firms (Porter, 1998). Thus, based on this reasoning the following hypothesis shall be tested in this research:

H2a: The financial performance of incubator production companies, which reside in a relatively more clustered region of within-industry localization, is comparatively better than those of isolated (less localized) incubator firms in the same industry.

- ➔ Hence, localized (within-industry clustered) incubators (financially) outperform less-localized incubators.

Nowadays, the formation of new businesses has also become an increasingly discussed topic. One of the key determinants of a successful spinoff of a new business comes down to the selection of an appropriate setup location. According to Schemmer (1978, 1982), it has been outlined that factors such as population growth (or/and density), availability of specialized personnel, including determinants such as wages and costs, could all be important drivers that lead towards agglomeration in specific areas. Furthermore, a positive and significant relationship between the presence of nearby rivals (in the same industry) and the likelihood of setting up a new business in the already clustered area has been uncovered (Dahl et al., 2009). In order to test whether investment in already clustered regions is more profitable than in isolated ones, the relation between the financial performance of isolated versus clustered new entrants needs to be investigated. Once again, it could be argued that due to already established clustered networks, firms are not only more likely to invest in such areas but are also expected to benefit more from engaging in such an action. Therefore, it could be reasoned that entrants should perform better due to the fact that they are to experience faster growth (better performance and greater profitability), stemming from the positive externalities of the industry cluster. Naturally, this leads to the next hypothesis tested in this paper:

H2b: The financial performance of new entrant production companies, which reside in a relatively more clustered region of within-industry localization, is comparatively better than those of isolated (less localized) new entrants in the same industry.

➔ Hence, localized (within-industry clustered) new entrants (financially) outperform less-localized new entrants.

Leaning on the above argumentation, another important factor of the firm's financial performance comes down to the internal economies of scale. In other words, it is also important to test whether the size of the firm matters when it comes down to profitability. Theory suggests that larger firms tend to generate superior profits due to the capacity or the possibility to experience internal economies of scale, mainly caused by the larger and more efficient usage of internal resources (Marshall, 1890).

H3: The financial performance of relatively large production companies is comparatively better than those of smaller firms in the same industry.

➔ Hence, large (in size) production firms outperform smaller organizations, due to internal economies of scale.

Having outlined the above-presented hypothesis, importantly, the following section will expound the data sources by providing a general overview of the key variables, tailed by a careful mapping of the empirical strategy which is used for the purpose of this research.

3. Data Sources and Empirical Strategy

Evidently, multiple research on industry clusters has been carried out for different markets around the globe. Yet, up to date and to the best of the author's knowledge, similar cluster examination is non-existent for the case of the Bulgarian market. The key reason for the lack of such important investigation can be credited to the scarcity of publicly available data. However, the author of this paper managed to attain practical records on registered companies from the pharma and oil production industries in Bulgaria for the period 2010-2016. The main source of information for this research comes from an online software of a regionally renowned business magazine Kapital.[4] Here it should be noted that to the author's best knowledge, the private organization Kapital holds the richest data of registered entities in Bulgaria. The records are obtained from: The National Register of Commerce, the National Statistical Institute, and the Bulgarian Stock Exchange. Combined, the whole dataset lists more than 200,000 organizations across Bulgaria. It is worth noting that for the purpose of this research, the author has manually collected all relevant figures.

Essentially, this paper tests the localization effect of the pharma and oil production companies only. To be precise, it is reported by Kapital (2018) that all pharmaceutical production companies (Table 1) are entities that are primarily engaged in: research, development, production, marketing, and sales of chemical or biological substances for medical or veterinary use (Kapital, 2018).[5] A brief overview of the relevant variables is situated in the summary statistics tables below.

[4] A full access to the website can be found at https://www.capital.bg/kpro/. However, only restricted access is available to non-registered accounts. For full access to the operational data, one should directly contact the author of this paper.

[5] Similarly, oil manufacturers (Table 2) primarily engage in: research, development, production, marketing and sales of crude oil and petrol substances for general use.

Table 1 - Summary Statistics for Pharma Industry

Variable	Obs.	Dummy or Discrete		Min	Max	Variable	Obs.	Continuous	Min	Max
		Median	Prop. (%)					Mean		
Year·	359	2013	-	2010	2016	**LQ Cluster**	359	1.340	0.0579	1.712
Entrant*	359	-	2.79	0	1	**Sales** ∟	359	17,510	0	318,854
Driven out*	359	-	2.23	0	1	**Profit** ∟	359	1,841	-2,197	40,885
Companies	62	-	-	50	54	**Tot Assets** ∟	349	31,217	0	586,509
						Tot Expenses ∟	359	13,163	0	272,171
						Tot Employees	359	143	0	2,020

∟ *Continues variables which are measured in thousands of BGN (Bulgarian currency). Data is collected at the end of the financial year for the forgone one.*

** Binary variables (1 coded when a company is reported to enter (registered) / exit (not registered) the industry in a given year, 0 otherwise). This variable ignores entrants/exits in the years 2010 and 2016 due to inability to follow up whether the company is registered (or not) in these specifics years (i.e. cannot be traced back due to missing information).*

∴ Obs. is the number of different companies in the whole sample for all years. The average (per year) number of tested companies is 51.

Gazing at the summary statistics tables (Pharma - Table (1) and Oil - Table (2)), the dependent variable takes the form of firm's financial figures (profit/loss balance or turnover) and serves as a proxy for firm's performance. The author has decided to use both figures for the purpose of this investigation, as this would give a better overview of the effect of the localization. It could be that localization variable is correlated with either of the indicators, only one (or none) of them. Yet, although the total number of observations for the pharma (oil) industry for the 7 years of reported data sums up to 359 (91) observations, the average (per year) number of companies is 51 (13). The small number of observations, specifically for the oil production sample, could be a reason for a limitation of the research. However, given the nature of the targeted companies, it is understandable that these industries are relatively less represented (in Bulgaria). In fact, such industries are characterized by small in number but larger in size entities. Nevertheless, the dataset from all years allows for a robustness checks in the form of a panel data investigation.

The key explanatory variable is measured by an LQ variable (localization employment quotient) and takes a "factor form" (more information can be found below and in the appendix B). This variable is measured in a specific way, traditional for calculating agglomeration economies

(Kukalis, 2010).[6] In theory, there are different models that calculate the density of a given industry (Kukalis, 2010). Due to the unavailability of other key data, the author of this paper took the decision to rely on this localization proxy. Moreover, the data is mathematically derived by using records from Kapital and NSI. The former is to calculate the employment figures, segregated by the different administrative geolocations of the pharma (oil) production factories, whereas the last is used to calculate the total employment of all industries and across all (six) regions in Bulgaria.

Table 2 - Summary Statistics for Oil Industry

Variable	Obs.	Dummy or Discrete		Min	Max	Variable	Obs.	Continuous	Min	Max
		Median	Prop.					Mean		
Year·	91	2013	-	2010	2016	**LQ Cluster**	91	1.049	0	5.673
Entrant*	91	-	6.59	0	1	**Sales** ∟	91	539,614	0	8,500,251
Driven out*	91	-	8.79	0	1	**Profit** ∟	91	-8,642	-532,347	118,891
Companies∴	19	-	-	10	16	**Tot Assets** ∟	85	350,233	0	4,532,168
						Tot Expenses ∟	91	516,098	0	8,209,312
						Tot Employees	91	179	0	2,223

∟ *Continues variables and are measured in thousands of BGN (Bulgarian currency). Data is collected at the end of the financial year for the forgone one.*

** Binary variables (1 coded when a company is reported to enter (registered) / exit (not registered) the industry in a given year, 0 otherwise). This variable ignores entrants/exits in the years 2010/2016 due to inability to follow up whether the company is registered (or not) in these specifics years (i.e. cannot be traced back due to missing information).*

∴ Obs. is the number of different companies in the whole sample for all years. The average (per year) number of tested companies is 13.

Other important control variables which can affect firm's performance include: firm size (i.e. total number of employees in the given year), total annual expenditure (i.e. total annual expenses that are spent on human capital, R&D and any other expenses related outflows) and total annual assets (Kukalis, 2010).[7] As witnessed in the above tables, all these controls are included in the regression estimations.[8] Although both organizations in the pharma and the oil are manufacturing, there are some conspicuous differences between these industries. Amongst all, the first thing that

[6] The exact measurement and explanation of the LQ variable are available in the next section and in Appendix B.

[7] All variables are the reported year-end ones, whereas the number of employees is the monthly average over the operating year.

[8] For a complete variable correlation matrix refer to Appendix D (for the Pharma sample) and to Appendix E (for the Oil sample).

comes to sight is that the mean profit for the pharma industry is positive, whereas a negative figure is reported for the oil industry.

As already stated, the aim of this paper is to identify the association between firm's localization and their performance, that is - the effect of being close to competition on the individual financial results of the organizations. In order to test whether agglomeration economies affect the financial performance of firms, first, it is important to test whether in fact there is a localization and to what extent it occurs.

3.1. Calculation of Localization Employment Quotient

In order to test the localization factor of the pharma (and oil) industry, the author of this paper leans on a localization quotient employment variable, which shows on average, a cluster factor (per region) for a specific industry in the tested country. In testing the extent of localization of the pharma (or oil) production industry across all 6 (six) regions in Bulgaria, the following localization quotient is used (Stavropoulos & Skuras, 2016):

$$LQ_{mj} = \frac{E_{mj}}{\sum_m E_{mj}} \Big/ \frac{\sum_j E_{mj}}{\sum_j \sum_m E_{mj}} \qquad (1)$$

Where Emj is the employment in region j and industry m, ∑m Emj is the total employment in all industries for region j, ∑j Emj is the employment for industry m in all regions, ∑j ∑m Emj is the total employment in all industries and all regions (Stavropoulos & Skuras, 2016). In essence, the larger the LQ factor, the greater the localization economies of the tested industry. The data is collected as follows: 1) for the full national and regional employment from NSI (National Stat. Institute); 2) for the pharma (oil) industry from the online database of Kapital Magazine.[9] In addition, there are six geographical regions in Bulgaria, across which the different companies are located. The real geolocation of the factory is used (as stated in Kapital) in order to correctly map the production factories in the corresponding regions (which is then used for the calculation of the employment in LQ above). Importantly, to fully understand the measurement of LQ and find out

[9] The full data of the regional employment (by cities) can be found in Appendix B under Figure (3) for pharma and Figure (4) for oil.

more information in regard to the general, administrative and demographic statistics, and the extent of localization and its measurement, one should carefully refer to Appendix (B).

Leaning on the above mathematical expression, it is possible to test how much on average, the pharma/oil industries tend to cluster based on the employment indicator and the region they operate in. Now, the usage of LQ might not fully depict the level of agglomeration. For instance, it could be that only one company resides in a given region but it does, however, employ a comparatively large proportion of the total industry employment (i.e. it could lead to overestimation).[10] However, the convenience of this model is that, it could be that smaller (in size) companies (i.e. fewer employees) can be situated in a highly localized region, hence the true benefits or the localization effects can be observed. In this way, the method allows testing between- and within-regional differences.

Now, in order to test the association between industry localization and the firm's financial performance, the above calculated LQ variable is inserted in the below-presented main regression estimation.

3.2. The Main Regression Estimation

The second step of the empirical model of this paper involves plotting the already found LQ variable (formula (1)) in the main regression formula (2). By comprising the above-described variables, the following specification model is employed:

$$\boldsymbol{Firm\ Performance_{ct}} = \boldsymbol{\beta 0} + \boldsymbol{\beta 1 LQ_{ct} x100} + \boldsymbol{\beta 2 Assets_{ct}} + \boldsymbol{\beta 3 Empl_{ct}} + \boldsymbol{\beta 4 Expenses_{ct}} + \boldsymbol{\varepsilon_{ct}} \quad (2)$$

$\boldsymbol{Firm\ Performance_{ct}}$ *is the dependent variable represented by total Turnover (Panel A (C) in Table 1 (2)) and Net Profit/Loss (Panel B (D) in Table 1 (2)) for company "c" at a time "t".*

$\boldsymbol{LQ_{ct}x100}$ *is a key explanatory variable and measures the localization quotient (i.e. proxy for clustering as explained above) for company "c" at a time "t". The variable is multiplied by factor of 100 (hundred) (and thus transformed into a percentage variable), thus allowing a greater variation and an easier interpretation of the main results.*

$\boldsymbol{Assets_{ct}}$ *is a control variable which measures the total level of assets of company "c" at a time "t".*

$\boldsymbol{Empl_{ct}}$ *is a control variable which measures the total number of employees of company "c" at a time "t".*

[10] Better explanation of the potential biasedness could be discovered in the limitations section.

$\boldsymbol{Expenses_{ct}}$ *is a control variable which measures the total level of expenses of company "c" at a time "t".*

$\boldsymbol{\varepsilon_{ic}}$ *is the error (idiosyncratic) term which includes residual variables (uncaptured by the specification) for company "c" at a time "t".*

Following the above regression model, the effect of clustering on firm's performance shall be studied by the $\boldsymbol{\beta 1}$ coefficient. In order to support the theoretical research, the expected sign of this coefficient is positive in direction. Taking into account the continuous nature of the dependent variable, the most appropriate statistical technique for this case is OLS.[11] The remainder of the Beta-coefficients study the relation between other associated variables that have proven their effect on firm's performance, namely firm's total assets, size/number of employees (H3), and total expenses.

[11] All regression calculations are carried out by the use of the statistical software "Stata". The full code of all estimations can be obtained upon request.

4. Analysis of the Main Results

Having understood the measurement of the key variables and the utilized empirical strategy, the paper can now focus on its major results. As described before, the key explanatory variable shows the average agglomeration across regions. Referring to Figures (3) and (4) in the appendix, it could be deduced that there are different levels of agglomeration across the six regions in Bulgaria. Essentially, the LQ variable measures how much the pharma (or oil) industry has a higher ratio of employment, compared to all other regions and at the same time taking into account the total level of country employment. Ultimately, the higher the LQ ratio, the higher the level of agglomeration for the tested industry in the given location. For instance, the LQ proxy of agglomeration of oil manufacturing firms in the Southeastern region for all years is above 5. This is to say that the oil manufacturing organizations in the Southeastern region of Bulgaria operate in 5 times (500% when inserted in the regression, see the estimation above) higher density region than the remainder of the regions (i.e. in a highly clustered area for this specific industry). It should be noted that the logic applies to all LQ measurements for both industries.[12]

The main results that delineate the association of pharma (oil) localization of companies with the firm's financial figures are plotted in Table 3 (Table 4) below. Conspicuously, the estimations are evident for the dependent variable of sales and profit, results for which are segregated in Panel A (C) and B (D) for the pharma (oil) industry, respectively.

Table (3) shows the results of various regression specifications for the pharmaceutical industry in Bulgaria. Regressions with a dependent variable of sales are situated in Panel A, whereas replicated models with a dependent variable of profit are located in Panel B.

Column (1) represents a naïve regression of the localization quotient on sales performance for all companies in the sample. Conspicuously, the beta-coefficient of the localization employment quotient (LQ) variable is not statistically significant, hence no initial correlation between localization economies with sales performance can be determined. That is to say that when holding everything else the same, being centered in a denser pharma area does not necessarily lead to greater annual sales figures.

[12] For a better and fuller explanation, please refer to Appendix (B).

Table 3 - Regression Results for Pharma Industry: Localization on Firm's Performance

	Dep. Variable: Annual Sales Panel A			Dep. Variable: Annual Profit Panel B		
	(1)	(2)	(3)	(4)	(5)	(6)
Specifications:	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**
LQ (H1)	**66.92**	**- .4556**	**- .7551**	**15.88***	**6.16***	**6.28***
	(58.25)	(16.04)	(16.02)	(6.61)	(2.31)	(2.54)
Sales					**.00269**	**.0086**
					(.0078)	(.0077)
Nr. of Empl. (H3)		**16.30** ˪	**17.07** ˪		**4.231***	**4.775***
		(7.88)	(7.85)		(1.144)	(1.267)
Expenses		**.5724***	**.5675***		**.0027**	**.0055**
		(.04261)	(.0423)		(.0076)	(.0072)
Tot. Assets		**.2519***	**.2515***		**.0418***	**.0352***
		(0228)	(.0227)		(.0038)	(.0041)
Entrant (~H2a)		**-1059.15**	**43.25**		**53.71**	**22.38**
		(4582.2)	(4521.23)		(661.28)	(636.35)
Driven Out (~H2b)		**199.4**	**-250.48**		**35.67**	**-104.03**
		(5427.6)	(5335.95)		(783.23)	(743.38)
Year Dummy	NO	YES	NO	NO	YES	NO
Panel Regr. (RE)	NO	NO	YES	NO	NO	YES
Nr. of Observations	**359**	**349**	**349**	**359**	**349**	**349**
Adj. R^2	0.001	.932	0.931	.0132	0.888	0.890

*Stat. sign at 98%; ˪Stat. sign at 95%; R^2 in column (3) and (6) is the overall one.

To further test the robustness of the initial findings, more controls are added in Column (2). The added control variables that associate with annual sales are number of employees, organization's total assets and total expenses (also entrant and driven out binary variables). To control for differences across years, a dummy variable is also included in the regression. Strikingly, although negative, the coefficient of the LQ variable remains economically insignificant, whereas the ones of the added variables (apart from entrant and driven out) are significant at very high levels.

To support the above findings, column (3) represents a panel data regression with random effects.[13] Similar to Column (2), the discoveries confirm the already imposed relationship between the LQ and the controls with the sales figures.

Overall, the full sample results for the pharmaceutical manufacturing industry shows no association between the annual sales dependent variable and localization of companies, hence the findings are in contradiction with the tested hypotheses. Evidently, the coefficient of the localization cluster shows a statistically insignificant relation, even once the controls are added (i.e. H1 cannot be confirmed). In addition, it is apparent that neither entrants (H2b) nor leavers (H2a) tend to be different from incubators. Based on the above, it could be deduced that all of the tested hypothesis (apart from H3) cannot be confirmed for the full data sample. In other words, there is no statistical significance between (more) clustered and isolated (less clustered) pharma manufacturers, when it comes to sales revenue.

Now, the author of this paper has decided to implement a further regression analysis. Based on the above argumentation, Panel B contains results for financial performance, measured by annual profit figures. Here, the regression model replicates the already presented path in panel A. Now, simply eyeballing the naïve regression in column (4), the explanatory variable is to be positive and highly statistically significant.

Yet, the reminder of the results in Panel B should be treated with caution. Unlike the findings reported in Panel A, adding further explanatory controls yields interesting discussion points. After adding the important determinants for sales, the main explanatory variable in column (5) remains highly statistically significant and positive in direction. In terms of magnitude, the coefficient predicts that on average, operating in a 100% denser (within-industry localized) area, on average leads to a more than 66,000 BGN increase in annual profit figures, when compared to

[13] A random effects model of panel data regression is used for all panel regressions. The reasoning behind this decision stems from additionally carried out testing (Hausman test located in Appendix C) and the fact that no systematic differences between companies have been identified. In other words, this test allows for between-companies investigation (however, no real within-company differences can be observed). Moreover, it should be stated that although the VIF and Correlation factors (some reported in appendix C) happen to suggest that there might be a high degree of correlation between the number of employees, total assets and total expenses. Nevertheless, it is also true that each of them, individually holds explanatory information. Supplementary testing with only one (or two) of the control variables (same specification models), still imposes similar results in regard to the LQ effect (R-squared still remains high). All it could be sensed here is that, in case of any biasedness, then we could be observing some degree of overestimation by the controls.

non-localized companies. Moreover, and perhaps as intriguingly, the additional controls remain positive and significant. In other words, having greater assets but also greater expenses is associated with larger profit figures for this particular industry. Importantly, having one more employee (Hypothesis 3) is associated with 4,231 BGN return on profits (on average) for the pharma production companies (column (5)). Additionally, the R-squared remains above 80% which means that these controls are rather imperative for the findings.[14] In addition, the robustness check carried out by the panel data regression model in Column (6) confirms the imposed findings.

Moreover, columns (5) and (6) further test H2a and H2b (i.e. new entrants and incubators performance). Notably, although all of the remainder of the coefficients remain relatively unchanged, we can see that the coefficients on entry and exit variables are statistically insignificant, hence do not bring any useful insights.[15]

Overall, the results in Table (1) show indication of a somewhat moderate association between the level of localization in the pharma industry and the firm's individual financial performance. It could be inferred that no clear difference between more and less clustered production companies can be seen when the dependent variable is proxied by the annual sales revenue accounts. Nevertheless, once the performance variable is measured by profit figures, there is a highly statistically significant connotation between the level of localization and the annual profits. In fact, being localized in a relatively denser area (of within pharma agglomeration) is, on average, associated with a significant increase in the profit numbers. Therefore, the more localized the area in which a pharma production company resides, regardless of its size, expenses and other controls, it still experiences economies from sitting in this cluster. Thus, stemming from the pharma sample results in Table (1), it can be concluded that agglomeration affects firm's profits in a positive way, however, no clear distinction for sales figures is to be observed, neither there is any relation with new entrants or industry-leavers.

[14] Of course, one should be cautious when interpreting the magnitude of the R-squared. It only suggests that the included control variables largely explain the difference of variance across the tested sample. Nevertheless, this is valid only if we can assure that the controls are not correlated with any other variables which remain in the error term.

[15] Additional testing with interaction terms confirm the same findings. The same is not included in the table, as the entrant and exit variable already impose no relation between the control and the dependent variable. Same applies to all similar cases.

Analogical to the above, Table (2) repeats the regression model, however, the tested industry in this case is the oil production firms in all regions in Bulgaria. Again, the dependent variable takes form of annual sales (Panel C) and annual profit figures (Panel D). An important notion here is that, unlike the pharma industry, the LQ variable has much greater standard deviation, which implies greater volatility of the levels of localization across regions.[16]

Focusing on the naïve regression (Column (1) in Panel (C) below) of the LQ variable on annual turnover, a positive and statistically significant coefficient can be identified. Importantly, once additional controls, such as firm's sales, size, total assets, and total expenses, are added to the regression (i.e. looking at Column (2)), the coefficient of the localization variable becomes insignificant. However, the importance of the control variables should not be neglected here. It has to be noted that the coefficients in front of the control variables (employees, expenses and total assets) are highly statistically significant. This means that there are other important factors that determine (or hold more information about) the annual sales of a given company. Hence, the underlying explanation suggests that sales figures are affected by other means, regardless of the level of localization, in which the company operates. Simply observing the coefficients of the added controls, it is important to note that all controls: total employees (H3), total assets and total expenses are positively related to sales figures, whereby the effect is larger on the first and smaller, but still present, for the latter two. In other words, an additional employee, on average, is associated with about 231,000 BGN increase in sales (of course one should see how different departments interact and which is the most beneficial), whereas, the greater the expenses, the grater the sales too. Nevertheless, and quite logical for the oil industry, as the total assets (i.e. total production capacity) of the company increases, then the sales revenues also increase. Nonetheless, here it should be stressed that neither the effectiveness of the employees (i.e. experience in the industry, age, skillset etc.), nor the direct reasons for expenses (i.e. investment, direct suppliers expense, R&D or employees' expenses) are taken into account, hence the effect should be treated as informative only.

[16] More details could be found in Appendix B.

Table 4 - Regression Results for Oil Industry: Localization on Firm's Performance

	Dep. Variable: Annual Sales			**Dep. Variable: Annual Profit**		
	Panel C			**Panel D**		
	(1)	(2)	(3)	(4)	(5)	(6)
Specifications:	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**	**OLS**
LQ (H1)	**7751.9***	**31.70**	**42.63**	**- 148.77***	**- 16.30**	**- 22.05**
	(767.1)	(63.61)	(61.64)	(35.41)	(35.41)	(34.43)
Sales					**.6354***	**.6401***
					(.0654)	(.0630)
Nr. of Empl. (H3)		**231.32***	**238.36***		**- 85.97***	**- 90.61***
		(54.93)	(53.981)		(34.07)	(33.60)
Expenses		**.9156***	**.9136***		**- .6172***	**- .6204***
		(.0163)	(.0158)		(.0606)	(.0582)
Tot. Assets		**.1049***	**.1034***		**-.0675***	**- .0674***
		(.0210)	(.0199)		(.0136)	(.0129)
Entrant (~H2a)		**9463.31**	**8722.12**		**-3338.7**	**- 4567**
		(2796)	(2726.09)		(15555.5)	(15196)
Driven Out (~H2b)		**-11321**	**6054**		**716.63**	**- 3226.8**
		(30984)	(26294)		(17234.8)	(14644)
Year Dummy	NO	YES	NO		YES	NO
Panel Data Regr.	NO	NO	YES		NO	YES
Nr. of Observations	**91**	**85**	**85**	**91**	**85**	**85**
Adj. R^2	0.529	0.998	0.973	0.156	0.699	0.623

*Stat. sign at 98%; ⌊Stat. sign at 95%; Standard errors are in parentheses; R2 in column (3) and (6) is the overall one.

Panel D uses annual profit figures as a dependent variable. Here the author tests the relation between LQ and annual profits in the oil industry. The naïve regression in column (4) identifies a negative and statically significant correlation between the localization quotient and the annual profit figures. However (and unlike the findings for the pharma industry), once the controls are added (i.e. column (5), the stat. sign. tendency seizes here. The full control regression shows that localization economies do not affect the profits of companies which operate in the oil industry. Once again, it becomes clear that there are more important factors that compose the financial performance of the company. Moreover, although all of the controls are statistically significant, the direction and the magnitude of the effects are different. Now, the findings in the oil sample

suggest that the annual sales are positively associated with the profitably of the company (which is normal and also seen in the pharma case). However, unlike the pharma data, the coefficient of the number of employees, expenses and total assets are negative in direction with the strongest effect stemming from the number of employees. Evidently, hiring one more employee could lead (on average) to 90,607 BGN reduction in profit figures. Given the market specificities of oil production industry, surely, hiring an additional employee could be linked with some form of direct (but also indirect costs). For instance, it could be that once the company decides to hire one extra person, there are ongoing indirect costs associated with this decision. Such decisions could be related to the following aspects: new machinery, higher transportation costs, greater inventory, hence lower profits in the short run.

Gazing at the robustness checks (column (6)), the panel data regression confirms the main findings from the pooled OLS in column (5). In reality, there are only minor changes in the magnitude of the effects, although the direction of the effects remains the same.

To further test H2a and H2b, the additional incumbents and leavers variables impose no significant effects, consequently no relation is to be seen when it comes to profitability. This is confirmed both in the findings stated in column (5) and in the robustness checks in column (6).

Stemming from the above, it could be construed that there are is no statistically proven relation or consequence from clustering in regions when it comes to the financial performance of the oil production companies in Bulgaria. It could be reasoned that there are market specific factors that correspond to these findings (same could be found in the discussion section).

Now, summing up the tendencies discovered in Table (1) and Table (2), intriguingly, it could be supposed that annual profits can have some dependent effect from the specific industry clustering, however, turnover is not really affected from this change. Notably, there are other important determinants for the financial prosperity of the organizations, with leading effects accredited to firm's total assets, number of employees (H3), and total expenses, whose direction could vary, depending on the market specificities of the tested industry.

5. Discussion of the Main Findings

In essence, the results of this research are informative about the relation between industry localization and the financial performance of the pharmaceutical (oil) production companies in Bulgaria for the period 2010-2016. Importantly, the main discoveries are quite distinguishable, not only within and between the tested industries, but also on which (financial) performance indicators the model is regressed. The fact of the matter is that the higher the localization factor, the greater the profitability of the pharma organizations (i.e. the more rivals stagnate within the same area, the greater the profit figures of the existing firms (on the individual level)). This association remains evident across all robustness checks. On the other hand, the statistical analysis inclines that there is no evidence of informative differences between clustered and isolated companies when their sales figures are tested as dependents. Summing both statements up, it turns out that localization of pharmacy production firms towards the competition could yield greater profit figures, however, this will not be reflected in the final sales statistics.

Intriguingly, for the oil industry, no real connection between the level of localization and the financial performance of the organizations is to be observed. In addition, the findings for both industries imply that factors such as number of employees, total expenses and total assets hold sufficient information which relates to financial performance, regardless of the dependent variable.

Yet, considering these bipolar findings, it is natural to dig in deeper to answer why such differences are observed and how they can be explained. Considering the scope of this investigation, it is quite challenging to narrow down to the exact reasons why the above phenomenon exists. First, the easiest way to capture a reasonable explanation is to understand the nature of the pharma (oil) industry in Bulgaria. Although both of the tested industries operate in the production field, there are market-specificities which drive the market supply and demand (and hence operational needs) in a fundamentally different way.

Most of the pharma production industry in Bulgaria are characterized by many (in quantity) but smaller (in size, i.e. number of employees) firms, which mainly serve the B2B market.[17] Comparatively, given the number of industry rivals (as seen in the summary stats), it could be

[17] The mean size variable is roughly the same for both industries. Moreover, production companies specialize in product development and production itself. The ready product is then sold to retailers (e.g. pharmacy (pharma) and petrol stations (oil)), before it reaches the final consumers (Kapital, 2018).

inferred that the pharma industry experiences fairer market competition, whereas the oil one leans towards a more monopolistic market. Since the location of the plants usually plays a crucial role for the development of the organizations in the corresponding industry, let us investigate both industries, one at a time.

5.1. Pharmaceutical Production Industry and Localization Economies Factor

The pharma production industry is characterized by various players in different regions in Bulgaria. The demand for Bulgarian production of drugs is increasing, not only locally, but also in the EU and other international markets (Kapital, 2018; NSI, 2018; Eurostat, 2018).[18] Therefore, the Bulgarian pharma production companies show increasing returns, positive annual growth, and steady pace of development. (Kapital, 2018).

Figure 1 – Clustering of Employees in the Pharma Manufacturing Industry
Average headcount (by regions) for the period 2010-2016

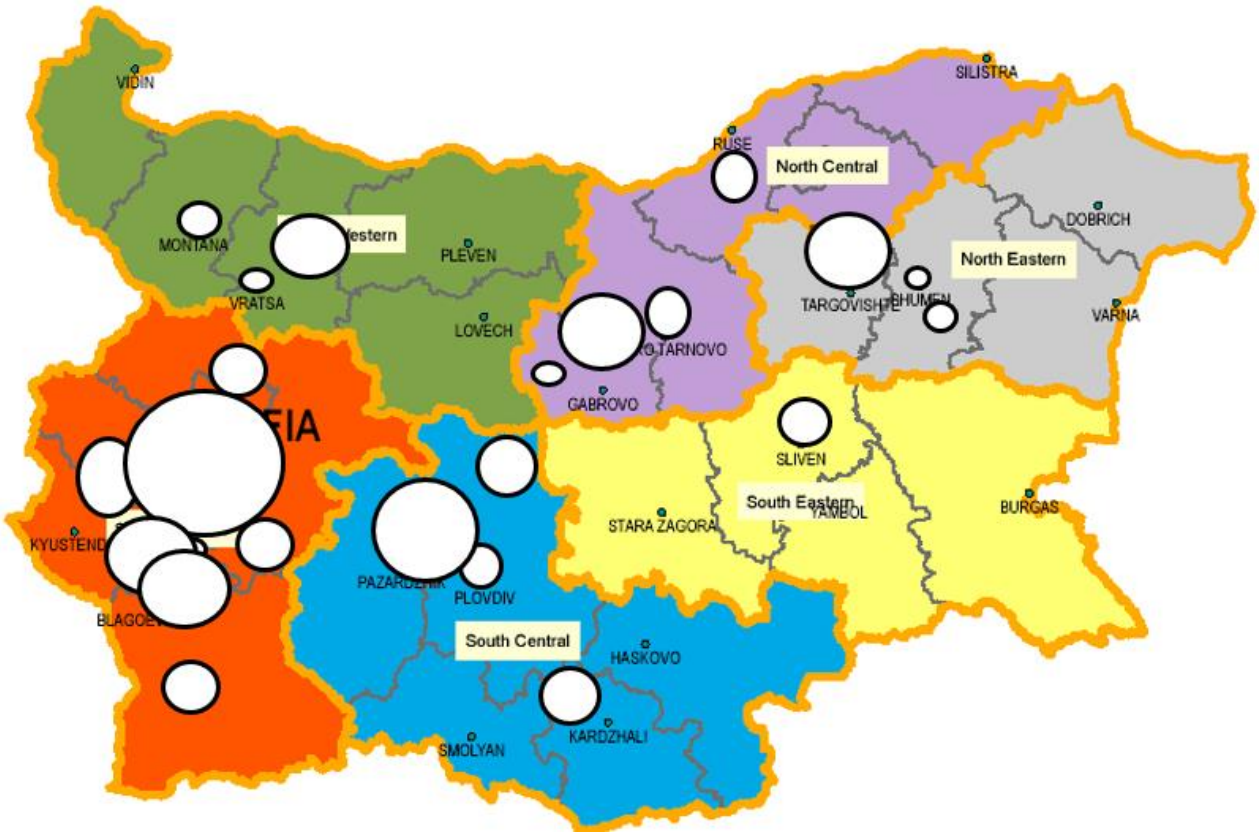


Figure (1) above represents an actual (geographical) mapping of the workforce, obtained from the total number of employees in the pharma sample (the greater the circle, the greater the number

[18] Refer to figure (17) in the appendix.

of employees in the area). Strikingly, it is evident that the pharma industry in Bulgaria (also confirmed by LQ variable) is relatively more clustered in specific areas. How this relates to the firm's performance is a different matter. Nevertheless, leaning on the results of this paper, it is evident that pharma production companies, in highly localized regions, do experience greater profits. Now, since there is no real difference in the annual sales of the production companies, then following the simple economics:

$$\textbf{Annual Profit} = \textbf{Annual Sales} - \textbf{Annual Expenses} \quad (3)$$

Now, ceteris paribus, the results of this study imply that greater levels of localization (i.e. higher LQ variable) in the pharma industry are associated with higher profits, but do not affect sales figures.

From the above, it follows that a greater localization factor (LQ) explains:

$$\textbf{LQ} \rightarrow \blacktriangle\textbf{Annual Profit} = \textbf{Annual Sales (const.)} - \blacktriangle\textbf{Annual Expenses} \quad (4)$$

Yet, for the annual profit to change, and given that the annual sales are not affected by the localization (i.e. they stay the same), then the difference should be equipoised by a change in the annual expenses. Since the LQ is positively related to annual profits, it follows that the annual expenses need to diminish, to offset the difference:

$$\textbf{LQ} (\uparrow) \rightarrow \blacktriangle\textbf{Annual Profit} (\uparrow) = \textbf{Annual Sales (const.)} - \blacktriangle\textbf{Annual Expenses} (\downarrow) \quad (5)$$

From the above, it follows that localizing in specific regions is related to either lower levels of total expenses or to greater external/internal economies of scale (i.e. increasing ROI ratios) stemming from sitting in a higher localized region. Importantly (to verify which tendency holds), it is also evident from the results that all expenses-related variables (i.e. hiring more employees, having greater total expenses and greater total assets) are positively related to profit figures. Combining the above, this is to say that the ROI (Return on Investment) or per employee (assets or expenses) cost is supposed to be lower for companies that localize in certain (e.g. more developed) geographical areas. Hence, localizing in such areas generates superior profits which makes companies to locate in such areas (this is how a cluster region is formed). It should, however, be noted that it is beyond the scope of this investigation to determine which exact economic factors within the specific areas affect the above-imposed tendency.

To dig in deeper, external economies of scale are reported to be caused by factors, depicted in the Porter's Diamond Model of Cluster economies (2000) and are as follows: Factor Conditions, Related Industries, Government and NGO Institutions, Supporting Industries, Rivalry, and Competitiveness across organizations. It could be that any of these factors actually lead towards greater annual profits.

In these lines, as the pharma production firms mainly sell in B2B markets, the supporting industries could play a crucial role when it comes to external economies of scale. To begin with, say we look at the logistics industry. Essentially, once the drugs are produced, inspected and packed, they need to be transported and delivered to domestic (or international) pharma retails. Therefore, the linkage with an adequate transportation system (i.e. highways, nearby airport, and railway) are of a high necessity. Once again, a quick look at figure (1) suggests that the pharma companies tend to cluster relatively more in the Southeastern (South-central) region in Bulgaria. It is a common fact that these are the better-established regions in the country with well-developed infrastructure and a nearby airport. These are all reasons that can serve as an explanation of external economies of scale (i.e. which lead to higher profits compared to less localized areas). Surely, depending on the operations of the company, due to improved logistics, firms can experience such economies of scale.

Yet, it is also important to reason why all control variables are positively related to the performance of the pharma production firms. Due to the nature of the business and the international patent requirements (GMP licenses), the only way for the pharma production companies to secure a long-term profitability is to carry out regular investments in R&D for new products (i.e. drugs) and/or to improve their patent defending mechanisms. Surely, such investments are heavy, however essential for the existence of the production firms. This reasoning serves as an explanation why the profitability of the pharma production companies is positively related to greater annual expenses (which include R&D and patents), greater assets (i.e. facilities or greater production) and hiring more employees. These suggestions are also confirmed by the discoveries in Table (1). Essentially, a newly patented drug (i.e. increased expenses) could logically be tailed by an upsurge in total assets (e.g. new production line), which in turn leads to an increase in the firm's size (i.e. hiring more employees). Logically, all these activities could together (or solely) lead towards increased annual sales and profit figures.

Leaning on the above argumentation, a future study should carry-out further investigation. For instance, checking whether R&D investment in year "t" is correlated with better financial performance in year "t+1", or "t+2". A future study could also focus more on the infrastructure factor. For example, including the number of main roads, overall traffic in the region and other important variables are good examples of logistics proxies.

5.2. Oil Production Industry and Localization Economies Factor

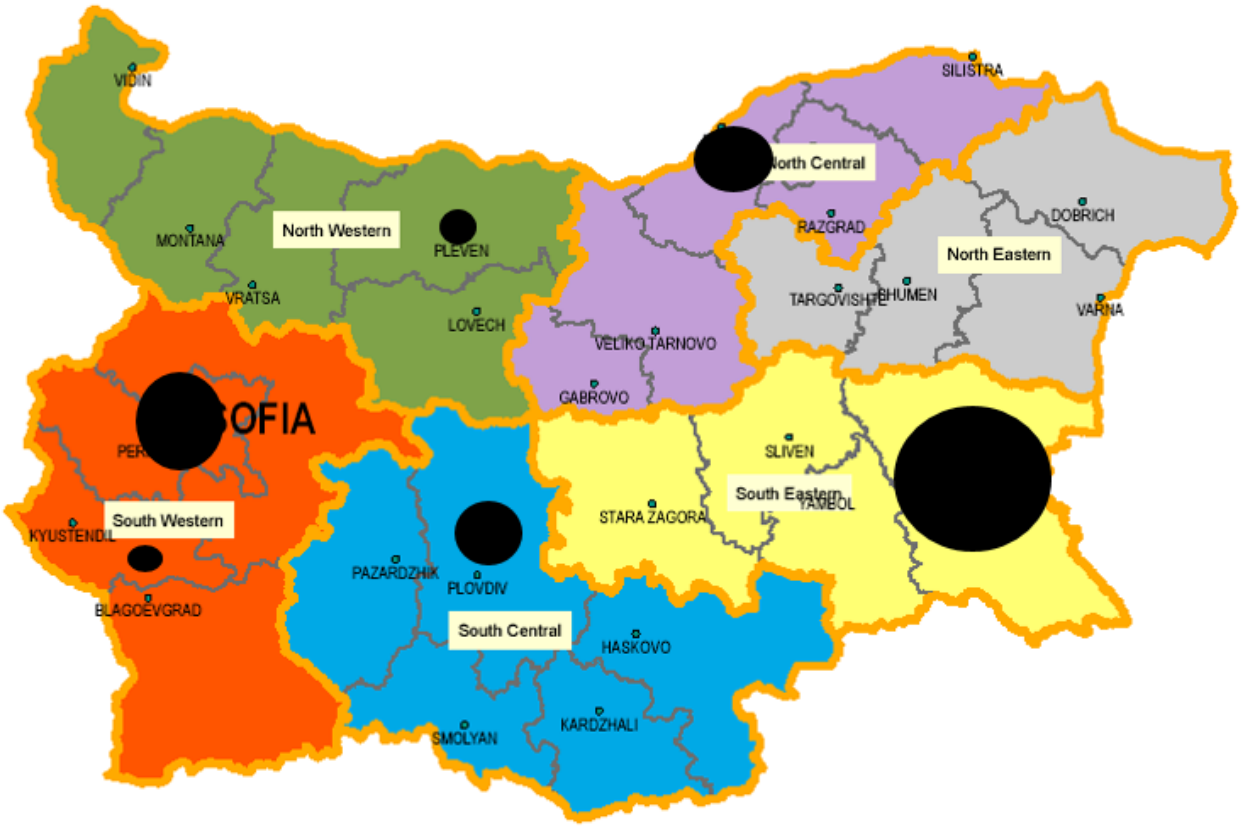
The oil production industry in Bulgaria is quite specific and is characterized by just a few market players, one of which is the largest (in net revenue) company (i.e. Lukoil) in Bulgaria (NSI, 2018; Kapital, 2018). In general, the market power is hugely diverted to this company and it would be fair to say that this industry features a monopolistic ecosystem, where the bargaining power is mainly in the hands of one company. Moreover, the market is inherently driven by the demand for oil products in the country. The entry barriers are also set high (e.g. expensive machinery, logistics channels, suppliers connections etc). Furthermore, the market is led by the "just enough" production capacity and the world trends in supply and pricing. Unlike the pharma case, the oil production companies mainly sell to domestic retailers, hence the demand is more or less predictable on year-to-year basis (Kapital, 2018).[19] Due to the nature of this particular business (i.e. the monopolistic setup and the fact that the industry operates in its sweet point or absolute scale), it is difficult for new entrants to exist in the long run.

Furthermore, as evident on the map below, the largest production company is based in the south-eastern region (very close to the Black Sea in the east). This fact makes sense since most of the crude oil supply is done by a sea transport, also oil extraction facilities are present as a supporting industry (Kapital, 2018). Nevertheless, as proven by the findings in this paper, even the higher clustering in this area does not lead to increased sales and/or profitability. This would imply that no localization economies are to be observed for this specific industry. As already argued, due to the monopolistic (at best oligopoly) nature of the oil production industry, the supply and demand are arguably regulated by the big players. Due to these reason, the localization economies could potentially be offset, thus no real positive externality can be derived from localizing in highly clustered regions. It should be pointed out that the small number of

[19] References for exports figures and internal market operations are available in appendix E (figure 17).

observations might also diminish the effect of the localization, hence testing larger geographical regions (e.g. Europe) could potentially show clearer estimates.

Figure 2 - Clustering of Employees in the Oil Manufacturing Industry
Average headcount (by regions) for the period 2010-2016



Essentially, the findings of this paper imply that other important factors such as the number of employees, annual expenses and total assets associate with sales and profit figures. Given the fact that the industry demand is predictable and does not vary by large, it means that an increase in the employees, assets or expenses would lead to a decrease in profitability. In other words, it would be difficult for incubators to sell more, given that there is no demand for this. If this happens, they should decrease prices in exchange for a greater demand. Nevertheless, the oil production companies have little (to no) incentive to do so (hence no greater supply is reasoned). The same is

also confirmed by the findings of this paper, where all controls are negatively related to profitability, i.e. the immediate effect from investment in them would decrease the total profits.[20]

To confirm these statements, a follow-up study needs to enlarge the geographical region of the oil industry or select a market with a more saturated presence of oil production companies. It also makes sense to consider the number of nearby companies (in the form of supporting industries), as well as the total number of suppliers.

[20] It should be noted that no long-term effects are observed. In other words, this paper only looks at the year-to-year change from these investments. To observe the big picture, a future study could also test how expenses in R&D in year "t" correlate with profit in year "t+1", for instance.

6. General Caveats and Room for Improvements

The main findings of this research shall be treated with caution, particularly when drawing general or causal conclusions. First, the author acknowledges that this thesis is a correlational study which does not necessarily convey causal interpretations. Yet, a general caveat of the paper might be the measurement of the localization variable (LQ) which can potentially be correlated with the number of employees' variable (size). The latter would be an issue (of a measurement error) when the number of companies within each region is insufficient and this leads to the LQ variable to be largely dependent on the actual number of employees in the given company in the corresponding region (as it is almost the same for the whole region). In particular, if this tendency holds, this might be observable in the oil production industry (mainly due to the small sample). Yet, another paper should try to involve different production companies or industries where the number of operating firms is larger. Moreover, this issue should not be problematic for the pharmaceutical industry, since in each of the tested region, there are companies with more (and/or less) than the national average number of employees for the pharma industry. In fact, this just brings further support to the key findings for the pharmaceutical production industry (i.e. there are companies within each region, which are different in size, however, still experience benefits from sitting in highly localized regions).

Another potential limitation of this paper comes down to the difficulty in conveying general conclusions. The research is based only on two production industries, where the oil one is much less represented in Bulgaria. Although, it manages to display a concrete picture for the Bulgarian production market, these findings might not necessarily follow the same pattern should the same be tested for other countries or regions, or even other production industries in Bulgaria. To fully answer the research question, a deeper investigation with greater number of production industries (and companies) needs to be carried out.

One could also question the fact that the market in Sofia (the capital) is far more developed than other cities in Bulgaria, therefore, the results from this research could be biased towards this region. However, as visible on Figure (2), not all production companies (i.e. oil for example)

localize within the Sofia region (Southwest).[21] Moreover, although the general overview is that companies tend to cluster towards different regions, it could be that companies stagnate more often in the larger (in population) cities, due to the generation of greater total demand (of course this depends on the industry itself). Nevertheless, as the LQ variable considers the total number of active labor, this somewhat controls for the number of citizens within the region.[22]

Finally, it should be argued that this paper should not suffer from reverse causality issues. It would be difficult to prove that profit (or sales) can lead to a specific localization of a single company towards other rivals. The profit of a single company depends on the performance of the firm itself and/or the performance of the whole industry (even if clustered). Nevertheless, the financial performance of a single firm can very rarely cause the same firm (or induce others) to stagnate, as this would require relocation, which, for production companies, is costly to be implemented. If anything, only new entrants might decide to locate in denser regions, however, for them no previous performance can be observed.

[21] According to statistics from the NSI, Magazine Kapital and Register of Commerce (2018), the automotive production companies in fact also localize in different areas other than the Southwest region. This of course is beyond the scope of this research but can be tested in a future paper.

[22] Yet, it should be noted that people might commute longer distances to reach their workplaces. This could place some thoughts whether LQ variable is correlated with the number of citizens. Furthermore, it could be that the number of working vs. total labor force vs. total population ratios vary across regions due to many reasons. Further statistics are present in Appendix A.

7. Recommendations for Future Research

Based on the above-presented limitations, a follow up study of the localization economy in Bulgaria should focus on two major things: First, an inclusion of richer dataset and second, try to identify the true factors that cause positive external economies of scale (if such are present).

Now, whereas the first one is simpler and can be carried out by getting the data, then the second one involves several steps. Following Porter's Diamond Model, one should carefully think of useful proxies for every single condition (e.g. a proxy for Supporting industries factor - one can use the number of companies within, say the logistics (or any other) industry; Government/State factor - one could use the total number of law changes that affect production capacity (or employees)); Factor conditions – e.g. the cost of squared meter; presence of natural endowments) and others. Once such regional factors are determined and collected, then these should be included in the main regression estimation to test their effects on performance.

Ideally, a more comprehensive paper should also use different measures of agglomeration economies, for example, the number of companies within specific diameter, or between-companies distance to confirm or reject the appropriateness of the LQ variable. Finally, a comparative research shall execute a similar approach to identify to what extent is there a localization of companies in other countries or markets (e.g. neighbor countries or EU) and how this affects the performance of the organizations. This way, it could be paralleled whether different countries should specialize in specific production (due to higher productivity) and use this comparative advantage to achieve greater overall economies of scale globally.

8. General Conclusion and Finale

With the ever-increasing speed of technological improvement, even more so of the information transmission, firms tend to specialize in specific goods or services, hence cluster together within specific areas, for yet to be verified reasons. Although sometimes this clustering can yield benefits, whereas in other occasions it can lead to drawbacks, it is impeccable to observe such stagnation, which would perhaps be even more profound with the future to come. There are numerous of successful examples of creation of globalized economies that in some cases become eco-systems. Economists argue that inter-company's specialization could benefit the nearby organizations, which in turn causes companies to experience externalities of scale. The same has been proven to be valid in various geographical regions, thus the author of this research decided to focus on the yet developing economies in Eastern Europe and Bulgaria in particular.

Yet, the importance of this study stems from the crucial role that the production companies used to play in Bulgaria during the USSR period. For many years (up to 1989) Bulgaria used to specialize in various production industries which involved heavy machinery etc. (Porter, HBS). Now, given the fact that Bulgaria is still considered to be a developing country within the EU, logically the income and the GDP per capita are still under the EU-average (Eurostat, 2018). Hence, due to obvious reasons there are many international investors which outsource their production to such locations. Since the production industry still proves its important role for the economy of Bulgaria, the author investigates whether localization of production companies is to be observed and if this stagnation leads to a better firm's performance. Indeed, the primary goal of this study is not to test the exact reasons why production companies cluster together, but to verify whether such stagnation leads to an improved firm's (financial) performance.

Leaning on the results and the discussion points of this research, the general conclusion implies that, within the tested period 2010-2016, the production companies do localize to different extent across the regions in Bulgaria. However, the financial consequences of doing so are somewhat differentiated between the tested pharma and oil industries. Evidently, in the pharmaceutical production industry, the more localized regions are associated with companies in generating higher profit figures (but no relation to the sales figures is identified). The same is mainly credited to the fact that pharma companies manage to generate positive externalities from the surrounding geographical area. On the contrary, the oil production findings suggest no association between

clustering and financial performance, which implies no real monetary incentive for companies to stagnate. It is also argued that the monopolistic nature of this industry dictates the market demand and "allows" no such externalities to be experienced. Furthermore, considering the additional testing, neither new entrants (start-ups), nor incubators are found to generate positive externalities from localizing in denser regions when it comes to the financial performance. Moreover, factors such as number of employees, total assets and expenses are also associated with changes (different in magnitude and in direction) in both profit and sales figures for both industries.

In summary, speaking from a purely financial perspective, firms could generate superior profits, which largely depends on the industry-specific cases. Although these results should not be generalized, they are somewhat instructive about the general overview of the production industries in Bulgaria for the period 2010-2016. Nevertheless, to prove the above statements, it is useful that a future study should include a greater scope of the tested production industries, add more controls, and distinguish between the factors that generate these external economies of scale (if such exist).

In a finale, keeping the pace of the technological improvement could essentially come down to the effective use of cluster economies, the consequence that it has on the surrounding businesses, and the surely upon the society. From a societal point of view, production companies would still continue to serve a driving role of delivering utmost goods, especially considering the fast pace of robotization and the ever-increasing sophistication of the consumer demand. Even though the study suffers from the above-described issues, yet it still points the essential need for further academic investigation in this subarea. Indeed, economic eco-systems take a central part in people's life and the overall economic development. Naturally, quarries that arise include: Is localization economies present and effective for specific industries only? Which are they? In what stage of the individual or collective development the clustering factor really matters? Do cluster economies lead to an improved performance and how this affects the surrounding environment? Surely, answers to these questions come up with effective and diligent research which requires great attention and support. Nevertheless, the practical understanding of how agglomeration economies develop is still a niche that remains to be discovered in more depth. Indeed, it is an important one that could lead to the refinement of the economic welfare of the society in general and thus deserves more profound attention, especially for the yet uninvestigated markets, just like the one in Bulgaria.

9. Appendices

9.1. Appendix A

Table 5 – GDP of Bulgaria in current prices (2010 – 2017) by NACE industry codes

Gross Domestic Product - (production approach)									
Current prices (in million BGN)									
Economic sectors and economic activity groupings	NACE Rev2 A10	**Gross Value Added-current prices**							
		2010	2011	2012	2013	2014	2015	2016	2017*
Agriculture, forestry and fishing	A	3110	3711	3741	3776	3819	3 664	3 817	3 693
Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply; water supply; sewerage, waste management and remediation activities	B_E	13149	16189	16370	16088	16484	18 013	19 813	20 606
Construction	F	4653	4497	4140	3431	3221	3 322	3 179	3 574
Wholesale and retail trade; repair of motor vehicles and motorcycles; transportation and storage; accommodation and food service activities	G_I	13107	14048	14040	15030	15486	16 884	17 970	18 511
Information and communication	J	3230	3776	3629	3987	4035	4 053	5 018	4 702
Financial and insurance activities	K	5892	5765	5414	4747	4977	5 355	5 867	6 394
Real estate activities	L	7853	8119	8177	8066	7489	7 535	7 777	8 764
Professional, scientific and technical activities; administrative and support service activities	M_N	3699	3836	4212	3980	4141	4 815	5 091	5 213
Public administration and defence; compulsory social security; education; human health and social work activities	O_Q	8880	9028	9569	9848	11079	11 037	10 837	12 003
Arts, entertainment and recreation, repair of household goods and other services	R_U	1451	1655	1697	1818	1879	1 868	1 849	1 954
Total Economic	**A+...+U**	**65022**	**70625**	**70989**	**70771**	**72610**	**76 546**	**81 218**	**85 413**
Adjustments (taxes less subsidies on products)		9749	10134	11052	11395	11024	12 025	12 912	13 218
Gross Domestic Product		**74771**	**80759**	**82040**	**82166**	**83634**	**88 571**	**94 130**	**98 631**
							National Statistical Institute, 2018		

9.2. Appendix B – Geographic specificities & LQ variable, its measurement and general information

Figure 3 – Administrative Map of Bulgaria (by Regions)

As shown on the figure above, Bulgaria is regionally divided in 6 (six) sub-areas (also referred to as NUTS 2 - Nomenclature des unités territoriales statistiques): North Western (BG31), North Central (BG32), North Eastern (BG33) and South Western (BG41), South Central (BG42), South Eastern (BG34) (Agency of Geodesy and Mapping, 2018). It should be clarified, that due to EU regulations, in 2007, Bulgaria is divided according to Eurostat nomenclature, which demands that the living population in each region must be between 800,000 – 3,000,000 number of people.

Collectively, the area (total of 110,995 sq. km), which each of these regions encompasses, is roughly the same as seen in the table on the right-hand side (Eurostat, 2018). This provides a reasonable ground for unbiased comparative statistics between the regions. Importantly, it should be noted that both the number of cities (villages) and the total population (hence working population and the number of active organizations) vary across regions. The same heterogeneity within regions can be observed in Figure (5) below, which shows the total employment statistics in Bulgaria for all NUTS-2 regions. Nevertheless, given the fact that the LQ proxy of localization takes into account both the regional and the national working population, it could be assumed that by large, these regional-specific differences are taken care of, hence no huge deviations (or biasedness) are to be expected in the regression estimations.

Figure 4 - Regional Land Coverage of Bulgaria (in sq. km)

GEO/TIME	2015
European Union	4 369 364
Bulgaria	110 995
North Western	19 068
North Central	14 812
North Eastern	14 647
South Eastern	19 801
South Western	20 300
South Central	22 366
UNIT	Square kilometre
LANDCOVER	Total land cover

Eurostat (2018)

Figure 5 – Total Employment Figures in Bulgaria for the period 2010 – 2016 (by regions)

Employment in Bulgaria (by years and regions)	2010	2011	2012	2013	2014	2015	2016
Total Employment	2242607	2242488	2218718	2226403	2240544	2254768	2277345
NorthWest	184320	185024	182472	182969	177648	175277	172591
NorthCentral	228917	227713	225124	225878	223150	222411	222962
NorthEast	263353	263168	262177	262269	263244	263035	262409
SouthEast	295672	294099	289827	290813	293783	295394	298121
SothWest	885419	887850	878376	885273	903080	914442	930842
SouthCentral	384926	384634	380742	379201	379639	384209	390420

NSI, 2018

Figure (5) shows the collective (and regional) employment figures of Bulgaria throughout the full 7 (seven) years of observations. The data is collected from the official website of the Bulgarian NSI (National Statistical Institute) and is freely accessible on the following web address: http://www.nsi.bg/en/content/13221/employment-and-hours-worked-regions.

It should be noted that the above figures are used for the calculation of the **LQ** variable. Namely: (1) for the $\sum_m$ **Emj** (the sum of the employment in all industries for region j) and (2) for the $\sum_j \sum_m$ **Emj** (the sum of employment in all industries and all regions). For instance, the total employment of the south-western region in 2016 is 930,842 number of employees (or $\sum_m$ **Emj**).

Then, the total national employment figure for 2016 is 2,277,345 number of employees across all regions and all industries (or $\sum j \sum m$ **Emj**).

Figure 6 – Regional Employment of the Pharma Manufacturers for the period 2010 - 2016

	Sum of Nr of Employees / Row Labels	Pharma Industry 2010	2011	2012	2013	2014	2015	2016	Grand Total
1.	**NorthWest Region**	246	261	231	152	130	120	117	1257
NW	Byala Slatina	203	215	178	101	80	68	65	910
NW	Montana	32	37	46	44	43	44	45	291
NW	Vratsa	11	9	7	7	7	8	7	56
	Sum of all ind. Northwest Regi	**184320**	**185024**	**182472**	**182969**	**177648**	**175277**	**172591**	**1260301**
LQmj	**Localisation Effect**	**0,4514**	**0,4572**	**0,3902**	**0,2490**	**0,2166**	**0,1985**	**0,1994**	**0,3089**
2.	**North Central Region**	392	383	409	450	453	434	425	2946
NC	Oreshak	8	10	10	11	13	11	11	74
NC	Ruse	21	20	23	27	23	1	0	115
NC	Troyan	339	318	336	362	364	366	358	2443
NC	Veliko Turnovo	24	35	40	50	53	56	56	314
	Sum of all ind. North Central R	**228917**	**227713**	**225124**	**225878**	**223150**	**222411**	**222962**	**1576155**
LQmj	**Localisation Effect**	**0,5791**	**0,5451**	**0,5600**	**0,5972**	**0,6009**	**0,5657**	**0,5607**	**0,5727**
3.	**North East Region**	**360**	**364**	**391**	**515**	**577**	**552**	**551**	**3310**
NE	Razgrad	327	341	377	492	553	525	527	3142
NE	Shumen	6	6	6	5	6	4	1	34
NE	Zavet	27	17	8	18	18	23	23	134
	Sum of all ind. North East Regi	**263353**	**263168**	**262177**	**262269**	**263244**	**263035**	**262409**	**1839655**
LQmj	**Localisation Effect**	**0,4623**	**0,4483**	**0,4597**	**0,5886**	**0,6488**	**0,6084**	**0,6177**	**0,5477**
4.	**South East Region**	**66**	**73**	**82**	**82**	**82**	**59**	**59**	**503**
SE	Sliven	66	73	82	82	82	59	59	503
	Sum of all ind. South East Regi	**295672**	**294099**	**289827**	**290813**	**293783**	**295394**	**298121**	**2057709**
LQmj	**Localisation Effect**	**0,0755**	**0,0804**	**0,0872**	**0,0845**	**0,0826**	**0,0579**	**0,0582**	**0,0752**
5.	**South West Region**	4240	4487	4730	4831	4906	5400	5331	**33925**
SW	Botevgrad	93	103	117	121	134	156	169	893
SW	Radomir	191	190	194	194	195	191	200	1355
SW	Sapareva Banya	29	27	26	27	26	26	25	186
SW	Sevlievo	5	4	4	4	4	4	4	29
SW	Sofia	3000	3154	3196	3152	3123	3390	3260	22275
SW	Sandanski	206	207	209	213	224	273	265	1597
SW	Dupnitsa	716	802	984	1120	1200	1360	1408	7590
	Sum of all ind. South West Re	**885419**	**887850**	**878376**	**885273**	**903080**	**914442**	**930842**	**6285282**
LQmj	**Localisation Effect**	**1,6195**	**1,6380**	**1,6599**	**1,6359**	**1,6081**	**1,7121**	**1,6846**	**1,6512**
6.	**South Central**	1327	1351	1355	1397	1421	1212	1259	9322
SC	Dimitrovgrad	47	38	30	31	27	28	27	228
SC	Peshtera	953	985	1007	1049	1078	1147	1198	7417
SC	Plovdiv	52	50	49	46	46	37	34	314
SC	Kazanluk	275	278	269	271	270	0	0	
	Sum of all ind. South Central R	**384926**	**384634**	**380742**	**379201**	**379639**	**384209**	**390420**	**2683771**
LQmj	**Localisation Effect**	**1,1659**	**1,1384**	**1,0970**	**1,1044**	**1,1080**	**0,9146**	**0,9486**	**1,0681**

The above data is collected from the annual (year-end) figures of the online magazine and the database of Kapital https://www.capital.bg/kpro/. As already explained, the gathered data is key for measuring the LQ localization variable (which serves as a proxy of localization economies) for

the pharma and oil industries in Bulgaria. The data in figures (6) and (7) measures: (1) the **Emj** which represents the employment stated for region j and industry m and (2) the **∑j Emj** which is the sum of employment for industry m in all regions.

Now, a caution should be given to the way in which the collected data is transformed into a useful tool to capture the effect of localization (i.e. measured by LQ). First, it should be noted that each organization (plant) is located in a specific city (or city area). Using this feature, it is possible to determine the exact geolocation of all operating plants (Head Quarters). Hence, as shown in figure (6) and (7), each region encompasses different cities. By grouping the employment in these cities (by regions), it is possible to determine both the national pharma (oil) employment and the total regional employment for the tested industries. For instance, looking at figure (6) of the pharma industry, the south-west region is represented by 7 (seven) cities in which companies within this industry operate. By adding the employment figures of these cities, we get a total of 5,331 workers (or **Emj)** in the south-west region in 2016. By summing the employment from all regions, we get a total of 7,742 workers (or **∑j Emj)** in the pharmaceutical industry (nationally) for 2016.

Now, by using the elements described above and relying on the below plotted formula, it is possible to figure out the LQmj for any of the 6 (six) regions (Stavropoulos & Skuras, 2016).

$$LQ_{mj} = \frac{E_{mj}}{\sum_m E_{mj}} \Big/ \frac{\sum_j E_{mj}}{\sum_j \sum_m E_{mj}}$$

For example, the LQmj of the SW region for the pharma industry in 2016 is 1.68. This means that on average, this region has 168% more employment for the pharma industry (when compared to the average regional pharma employment and controlling for the total national employment)). Once the LQmj variable is calculated, it is then inserted next to all companies, which reside in the corresponding geographic region for the given year. For examples: If company X resides in the south-west region, in year 2016 the corresponding LQ variable for company X is 1.68. The process is repeated for all manufacturers across all regions and years for the pharma (and the oil) industry. Moreover, gazing at the LQmj variables (in figure (6) above), it is evident that the south-west and the south-central regions have higher values of localization throughout the tested years. This would suggest that one should expect higher levels of clustering of companies that reside in these two areas, hence the positive/negative externalities, on average, should be more profound for companies which reside in these two areas (i.e. what is tested in this research). It should also be

noted that each of the reported cities has at least one pharma manufacturing company (unlike the oil industry, where regions with only 1 (one) or none represented companies is to be observed). This largely stems from the market specificities of the corresponding industry (more is explained in the discussion section of this paper).

Figure 7 – Regional Employment for the Oil Manufacturers for the period 2010 - 2016

	Sum of empl	Oil industry							
	Row Labels (years)	2010	2011	2012	2013	2014	2015	2016	Grand Total
1.	**South East**	2223	1885	1739	1578	1497	1447	1410	11779
SE	Burgas	2223	1885	1739	1578	1497	1447	1410	11779
SE	Tvurditsa	0	0	0	0	0	0	0	0
SE	Yambol	0	0	0	0	0	0	0	0
	Sum of all industries in SE	295672	294099	289827	290813	293783	295394	298121	
LQmj	**Localisation Effect**	5,6731	5,4417	5,4722	5,3904	5,2834	5,2142	5,1581	
2.	**South West**	478	482	436	423	420	428	439	3106
SW	Sofia	478	482	436	423	420	425	436	3100
SW	Kostinbrod	0	0	0	0	0	0	0	0
SW	Krupnik	0	0	0	0	0	3	3	6
	Sum of all industries in SW	885419	887850	878376	885273	903080	914442	930842	
LQmj	**Localisation Effect**	0,4074	0,4609	0,4527	0,4747	0,4822	0,4982	0,5143	
3.	**South Central**	104	113	101	96	97	97	96	704
SC	Rakovski	104	113	101	96	97	97	96	704
	Sum of all industries in SC	384926	384634	380742	379201	379639	384209	390420	
LQmj	**Localisation Effect**	0,2039	0,2494	0,2419	0,2515	0,2649	0,2687	0,2682	
4.	**North East**	108	111	110	101	103	101	97	731
NE	Ruse	108	111	110	101	103	101	97	731
	Sum of all industries in NE	263353	263168	262177	262269	263244	263035	262409	
LQmj	**Localisation Effect**	0,3094	0,3581	0,3826	0,3826	0,4057	0,4087	0,4031	
5.	**North Central**	0	0	0	0	0	0	0	
	Sum of all industries in NC	184320	185024	182472	182969	177648	175277	172591	
LQmj	**Localisation Effect**	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	
6.	**North West**	0	0	0	0	0	1	0	1
NW	Pleven	0	0	0	0	0	1	0	1
	Sum of all industries in NW	184320	185024	182472	182969	177648	175277	172591	
LQmj	**Localisation Effect**	0,0000	0,0000	0,0000	0,0000	0,0000	0,0061	0,0000	
	Grand Total	2913	2591	2386	2198	2117	2074	2042	16321

9.3. Appendix C – Hausman

Figure 8 - Pharma Industry - Hausman Test for Fixed or Random Effects

	Coefficients			
	(b) fixed	(B) .	(b−B) Difference	sqrt(diag(V_b−V_B)) S.E.
lq2	593.0773	628.5101	−35.43278	235.1145
sales	.0443377	.0086537	.035684	.002668
empl	−3.163965	4.775149	−7.939114	2.644502
exp	.0127315	.0055021	.0072293	.
assts	−.0254198	.0352387	−.0606585	.0043131
enter	186.7685	22.38436	164.3841	393.8715
exit	−112.901	−104.0395	−8.861558	.

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(4) = (b−B)'[(V_b−V_B)^(−1)](b−B)
= 9.38
Prob>chi2 = 0.0522

Gazing at the Hausman test for fixed and random effects, it can be reported that the H0 cannot be rejected. In other words, it cannot be concluded that fixed effects model is the most appropriate for this study. Furthermore, carrying out similar testing for all regressions for both industries yields similar results. Overall, it could be stated that no systematic differences in the error terms of the tested variables is to be observed. Therefore, it can be inferred that there are no systematic (i.e. within-company) differences across companies (and years) which correlate with the dependent variable. Although this might be difficult to believe, it could be that more controls should be inserted in the regression to account for this or the number of tested companies does not suffice. In other words, we cannot account for within-companies differences, hence we look at the between-companies (or regional) differences and use random effects model.

9.4. Appendix D – Pharma Sample (Correlation Matrix and Useful Graphs)

Figure 9 – Pharma Sample (Variables Correlation Table)

	yr	lqr	sales	profit	empl	asst	exp
yr	1.0000						
lqr	0.0057	1.0000					
sales	0.0532	0.0607	1.0000				
profit	0.0073	0.1262	0.8925	1.0000			
empl	0.0243	0.0437	0.9312	0.9113	1.0000		
asst	0.0242	0.1018	0.9322	0.9393	0.9398	1.0000	
exp	0.0319	0.0268	0.9227	0.8264	0.8970	0.8496	1.0000
enter	0.1195	0.1017	-0.0570	-0.0521	-0.0729	-0.0549	-0.0565
exit	0.0858	0.0208	-0.0436	-0.0440	-0.0498	-0.0399	-0.0419

	enter	exit
enter	1.0000	
exit	0.0891	1.0000

Figure 10 - Pharma Sample – Variable Correlation Matrix

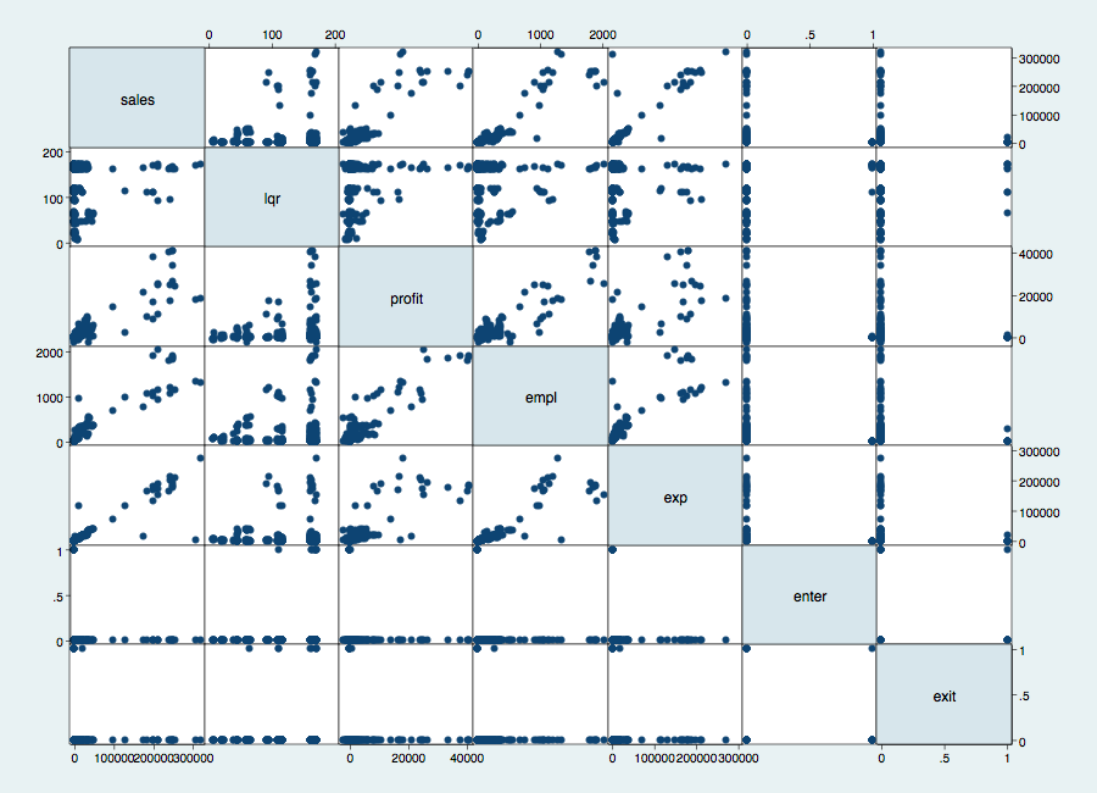


Figure 11 - Pharma - 3D mapping of profit (y-axis), LQx100 (x-axis) and number of employees (`size)

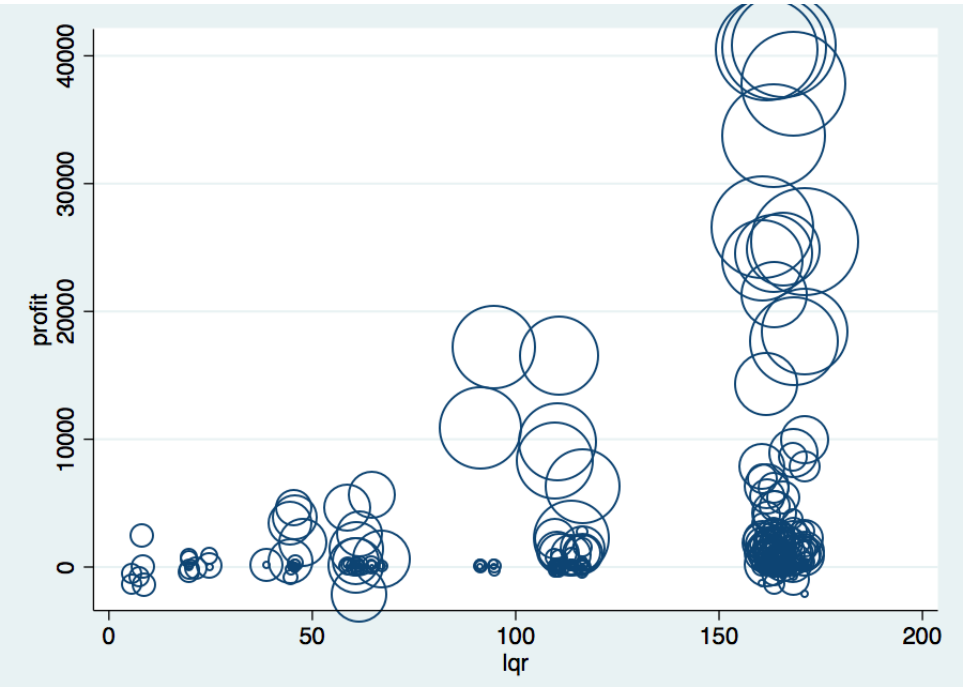
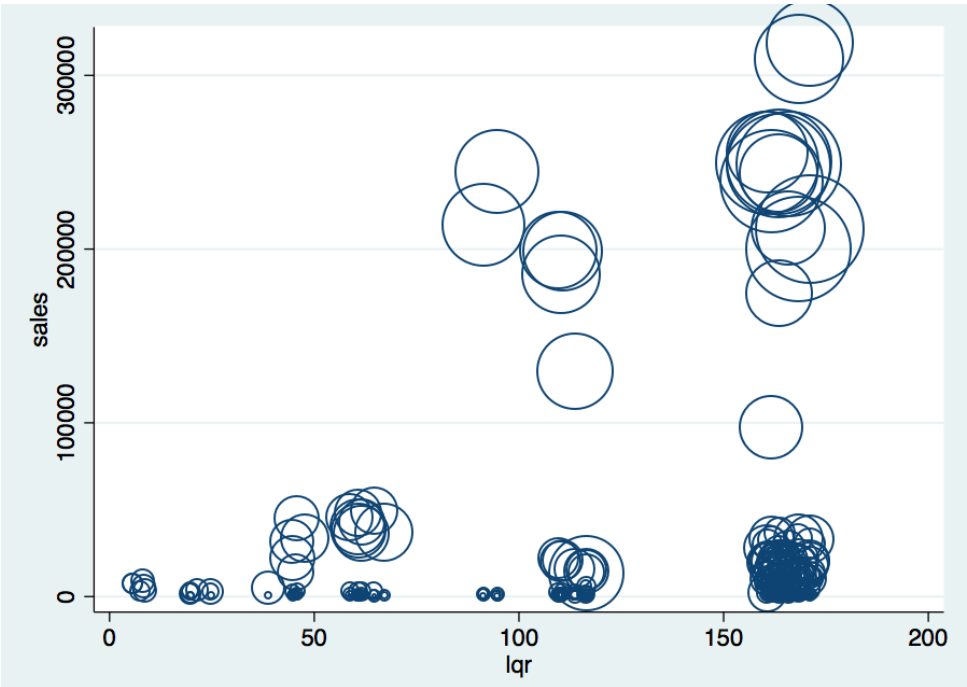


Figure 12- Pharma - 3D mapping of Sales (y-axis), LQx100 (x-axis) and number of employees (circle size)



9.5. Appendix E – OIL Sample (Correlation Matrix and Useful Graphs)

Figure 13 - Oil Sample (Variables Correlation Table)

	yr	lqr	sales	profit	empl	asst	exp
yr	1.0000						
lqr	-0.0901	1.0000					
sales	0.0080	0.7309	1.0000				
profit	0.0100	-0.4069	-0.5513	1.0000			
empl	-0.0040	0.7070	0.9628	-0.5135	1.0000		
asst	0.0900	0.7168	0.9485	-0.5317	0.9238	1.0000	
exp	0.0028	0.7271	0.9990	-0.5747	0.9575	0.9412	1.0000
entrant	0.1381	0.0252	-0.0806	0.0373	-0.1064	-0.0914	-0.0798
exit	-0.0046	0.0922	-0.0942	0.0434	-0.1245	-0.0991	-0.0932

	entrant	exit
entrant	1.0000	
exit	0.0739	1.0000

Figure 14 - Pharma Sample – Variable Correlation Matrix

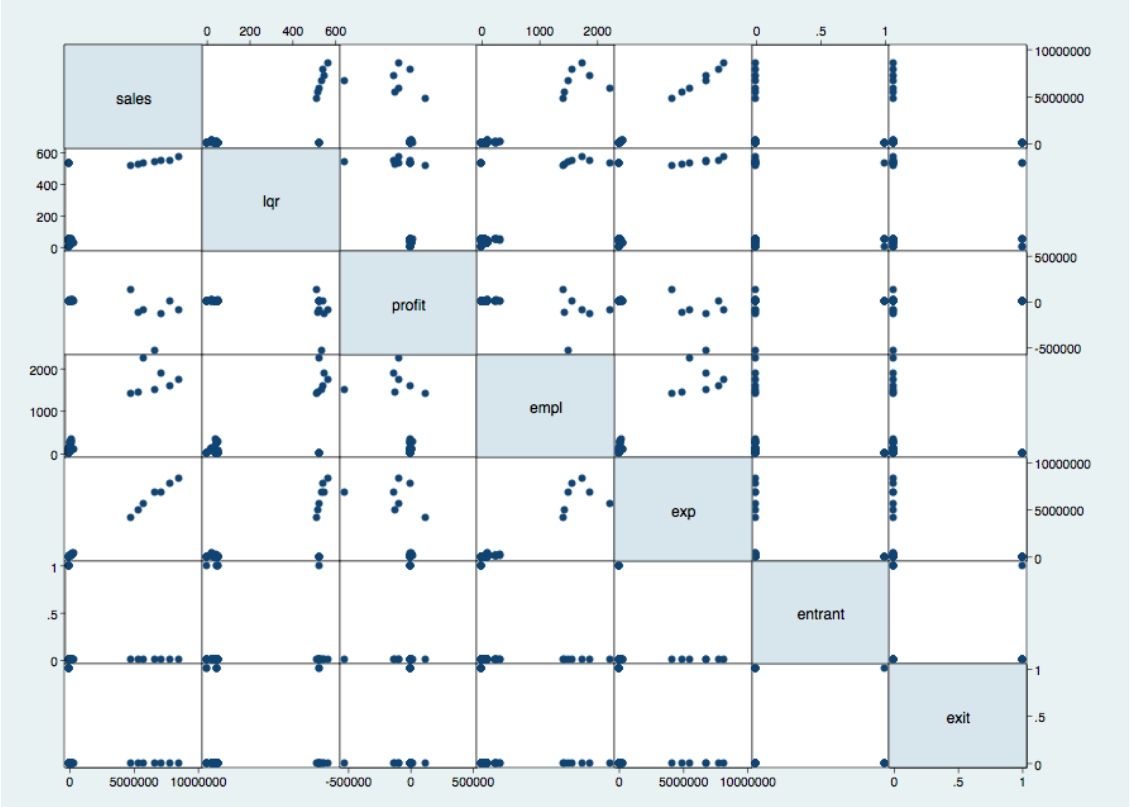


Figure 15 - Oil - 3D mapping of profit (y-axis), LQx100 (x-axis) and number of employees (circle size)

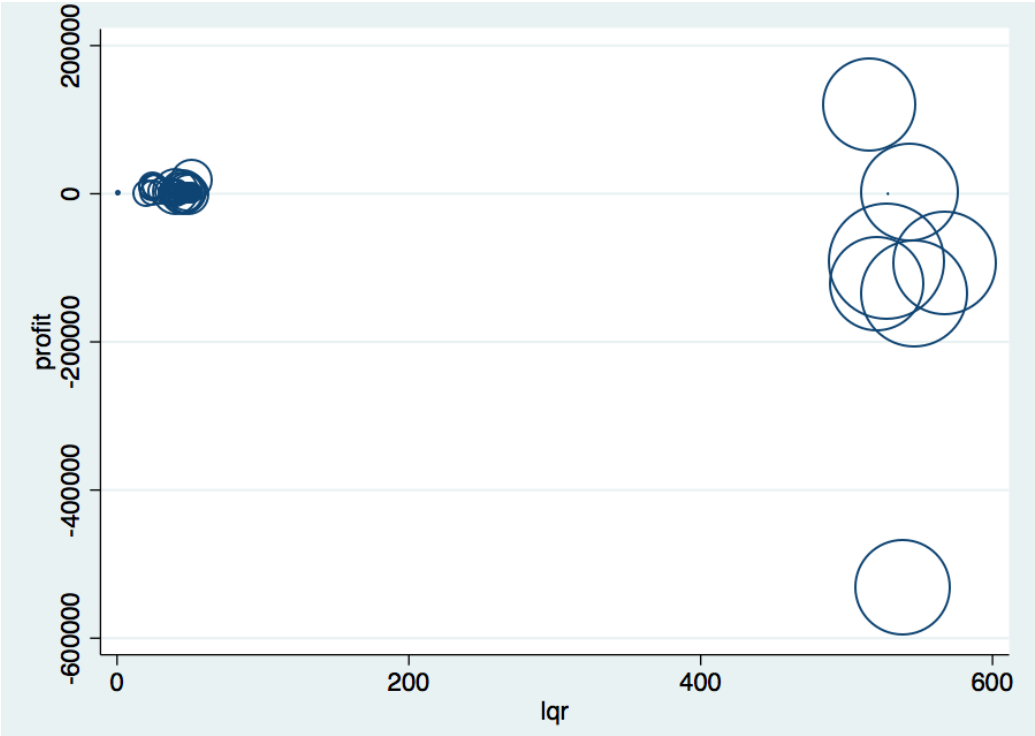


Figure 16 - Pharma - 3D mapping of Sales (y-axis), LQx100 (x-axis) and number of employees (circle size)

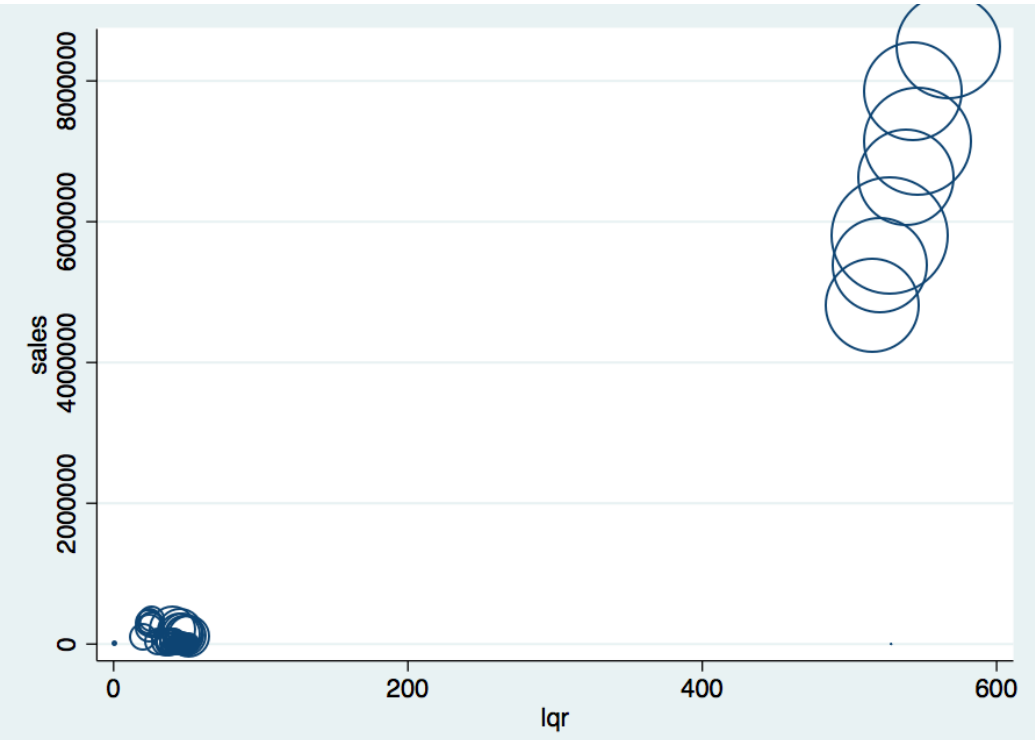


Figure 17 - Imports and Exports for Pharma and Oil related industries

Sections by SITC	Trade balance - FOB/CIF	Average year ratio
	Million BGN Trade Balance for the period 2010 - 2016	Change of Exports-to-Imports Ratio for 2010-2016
Mineral fuel, lubricants and related materials	-30058.2	0.174
Chemical and related products n.e.c.	-17190.5	1.301

Source: NSI, 2018

The above table shows the trade balance by SITC sectors in Bulgaria for the period 2010 – 2016. From the figure, it is evident that the total trade deficit for mineral fuel (includes oil production) is almost double than the trade deficit for chemical and related products (includes pharma production). Yet, since the total traded goods for the oil industry (in relative terms) outnumber the ones in the pharmaceutical industry, this information might not always be the best factor to look at. Thus, it would be good to look at the very right column, which depicts the exports-to-imports growth ratio (as an average) for the period 2010-2016. Gazing at the oil parameter 0.174), it means that the average yearly growth change of imports exceeds that of the exports, hence a negative long-term trend can be imposed (i.e. decrease in exports). On the other side, the ratio of 1.301 for the pharma industry, suggests the average yearly growth change of exports exceeds that of the imports, hence a positive long-term trend exists (i.e. increase in exports).

10. Reference List

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science*, *45*(7), 905-917.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, *86*(3), 630-640.

Austin, S. B., Melly, S. J., Sanchez, B. N., Patel, A., Buka, S., & Gortmaker, S. L. (2005). Clustering of fast-food restaurants around schools: a novel application of spatial statistics to the study of food environments. American journal of public health, 95(9), 1575-1581.

Bell, C., & Lewis, M. (2005). Economic Implications of Epidemics Old and New.

Burt, M. G. (1992). The justification for applying the effective-mass approximation to microstructures. *Journal of Physics: Condensed Matter*, *4*(32), 6651.

Canina, L., Enz, C. A., & Harrison, J. S. (2005). Agglomeration efects and strategic orientations: Evidence from the US lodging industry. *Academy of management journal*, *48*(4), 565-581.

Ciccone, A., & Hall, R. E. (1993). *Productivity and the density of economic activity* (No. w4313). National Bureau of Economic Research.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, *94*, S95-S120.

Dahl, M. S., & Sorenson, O. (2009). The embedded entrepreneur. *European Management Review*, *6*(3), 172-181.

Enright, R. D. (1996). Counseling within the forgiveness triad: On forgiving, receiving forgiveness, and self-forgiveness. *Counseling and values*, *40*(2), 107-126.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of sociology*, *91*(3), 481-510.

Harrison, A. E. (1994). Productivity, imperfect competition and trade reform: Theory and evidence. *Journal of international Economics*, *36*(1-2), 53-73.

Harrison, B., Kelley, M. R., & Gant, J. (1996). Specialization versus diversity in local economies: The implications for innovative private-sector behavior. *Cityscape*, 61-93.

Henderson, J. V. (1986). Efficiency of resource usage and city size. *Journal of Urban economics*, *19*(1), 47-70.

Henderson, V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic growth*, *8*(1), 47-71.

Holmes, S. (1996). *The anatomy of antiliberalism*. Harvard University Press.

Jacobs, J. (1969). Strategies for helping cities. *The American Economic Review*, *59*(4), 652-656.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, *108*(3), 577-598.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, *99*(3), 483-499.

Kukalis, S. (2010). Agglomeration economies and firm performance: the case of industry clusters. *Journal of Management*, *36*(2), 453-481.

Martin, R., & Sunley, P. (2003). Deconstructing clusters: chaotic concept or policy panacea?. *Journal of economic geography*, *3*(1), 5-35.

Marshall, G. (1980). *Presbyteries and profits: Calvinism and the development of capitalism in Scotland, 1560-1707*. Oxford University Press, USA.

Maskell, P. (2001). Towards a knowledge-based theory of the geographical cluster. *Industrial and corporate change*, *10*(4), 921-943.

Malecki, E. J. (1985). Industrial location and corporate organization in high technology industries. *Economic Geography*, *61*(4), 345-369.

Myles Shaver, J., & Flyer, F. (2000). Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic management journal*, *21*(12), 1175-1193.

Porter, Michael E. The Competitive Advantage of Nations. *New York: Free Press*, 1990.

Porter, M. E. (1998). *Clusters and the new economics of competition* (Vol. 76, No. 6, pp. 77-90). Boston: Harvard Business Review.

Porter, M. E. (2000). Location, competition, and economic development: Local clusters in a global economy. *Economic development quarterly*, *14*(1), 15-34.

Rosenfeld, S. A. (1997). Bringing business clusters into the mainstream of economic development. *European planning studies*, *5*(1), 3-23.

Saxenian, A. (1994). Regional networks: industrial adaptation in Silicon Valley and route 128.

Sørensen, J. B., & Sorenson, O. (2003). From conception to birth: Opportunity perception and resource mobilization in entrepreneurship. In *Geography and Strategy* (pp. 89-117). Emerald Group Publishing Limited.

Stavropoulos, S., & Skuras, D. (2016). Firm profitability and agglomeration economies: An elusive relationship. *Tijdschrift voor economische en sociale geografie*, *107*(1), 66-80.

Storper, M. (1989). The transition to flexible specialisation in the US film industry: external economies, the division of labour, and the crossing of industrial divides. *Cambridge journal of economics*, *13*(2), 273-305.

Tallman, M. S., Kim, H. T., Paietta, E., Bennett, J. M., Dewald, G., Cassileth, P. A., ... & Rowe, J. M. (2004). Acute monocytic leukemia (French-American-British classification M5) does not have a worse prognosis than other subtypes of acute myeloid leukemia: a report from the Eastern Cooperative Oncology Group. *Journal of clinical oncology*, *22*(7), 1276-1286.

Online References

A. (2018). Agency of Geodesy and Mapping. Retrieved August 2018, from http://www.cadastre.bg/about-us/

C. (2018). Класацията Капитал 100 – най-големите компании в България. Retrieved August, 2018, from https://www.capital.bg/k100-2018/

C. (2018). Капитал - прогнози, които се сбъдват. Новини и анализи от България, политика и икономика, финансова информация, начин на живот. Retrieved August, 2018, from https://www.capital.bg/

E. (2018). Your key to European statistics. Retrived June, 2018, from https://ec.europa.eu/eurostat/

N. (2018). National Statistical Institute. Retrieved July, 2018, from http://www.nsi.bg/en/