2008

AGGLOMERATION ECONOMIES AND GEOGRAPHIC CONCENTRATION OF MANUFACTURING IN UKRAINE

Volodymyr Vakhitov
University of Kentucky, vakhitov@eerc.kiev.ua

ABSTRACT OF DISSERTATION

Volodymyr Vakhitov

The Graduate School

University of Kentucky

2008

AGGLOMERATION ECONOMIES AND GEOGRAPHIC CONCENTRATION OF MANUFACTURING IN UKRAINE

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Business and Economics
at the University of Kentucky

By
Volodymyr Vakhitov

Lexington, Kentucky

Director: Dr. Christopher Bollinger, Professor of Economics

Lexington, Kentucky

2008

ABSTRACT OF DISSERTATION

AGGLOMERATION ECONOMIES
AND GEOGRAPHIC CONCENTRATION
OF MANUFACTURING IN UKRAINE

As a post-Soviet economy, Ukraine has inherited substantial production assets and qualified personnel. However, the economy was dominated by large-scale enterprises designed for much bigger markets. After the collapse of the Soviet Union Ukrainian firms faced lack of planning, breaks in contacts with their former suppliers and customers, and distortion of prices. There was a clear need in restructuring of the entire economy. Restructuring included splitting firms into smaller parts and privatization. The first phase of transition was completed by 2000 when the output grew for the first time after a long recession in nineties, and most firms became private property.

In this work I explore trends in geographic and industrial concentration of Ukrainian manufacturing firms over the period of 2001 to 2005. I found that this period was characterized by relocation of firms between sectors and between regions, as well as by an increase in economic concentration of industries. The speed of adjustment was different for various sectors and even for different industries within manufacturing. Even though the economy is still dominated by large firms, the average firm size decreases due to a rapid growth in the number of new firms. Geographically, manufacturing tends to increasingly concentrate mostly around a few big cities, apparently at the expense of other regions.

I also estimate the external scale effects and compare them with Western studies. In particular I focus on machinery and high tech. I found strong localization and urbanization effects in both industry groups. An important contribution of this work is the analysis of the effect of ownership structure on agglomeration economies. I found that private firms tend to enjoy external scale effects to a greater extent than state owned, and foreign owned firms appear to be the most efficient in extracting benefits form agglomeration.

Aggregation of the data may distort the estimates of agglomeration effects. I show that most effects take place at the nearest neighborhoods. When the physical distance between firms increases agglomeration effects attenuate quickly. However, localization effects reveal themselves at different level of industrial aggregation for various industries. This may reflect more complicated relationships within sectors and requires further analysis.

KEYWORDS: Agglomeration economies, geographic concentration, data aggregation, transition economies, ownership structure.

Volodymyr Vakhitov

July 23, 2008

AGGLOMERATION ECONOMIES
AND GEOGRAPHIC CONCENTRATION
OF MANUFACTURING IN UKRAINE

By

Volodymyr Vakhitov

Dr. Christopher Bollinger

Director of Dissertation

Dr. William Hoyt

Director of Graduate Studies

July 23, 2008

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors.

Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Volodymyr Vakhitov

The Graduate School

University of Kentucky

2008

AGGLOMERATION ECONOMIES AND GEOGRAPHIC CONCENTRATION OF MANUFACTURING IN UKRAINE

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Business and Economics at the University of Kentucky

By
Volodymyr Vakhitov

Lexington, Kentucky

Director: Dr. Christopher Bollinger, Professor of Economics

Lexington, Kentucky

2008

ACKNOWLEDGMENTS

"*It is impossible to have a child and two PhDs in the family at the same time.*"
Kenneth Gray, Director of Master Program in Economics, Economic Education and Research Consortium, Kyiv, Ukraine, 1998

The following dissertation, while an individual effort, would never have happened without help of many people. First of all I would like to thank to my thesis advisor, Chris Bollinger, for continuous support, patience and hours we spent discussing various aspects of this research. The several last months gave me more experience in the academic work than any coursework I took before. Thank you for leading me through the process, for making me believe that I could achieve this result, and for the superb guiding through the whole way. Thank you for not giving up on me!

I would also like to thank to my other committee members, Dr. Bill Hoyt, Dr. Ken Troske, Dr. Alison Davis, and the outside reader, Dr. Pradyumna Karan, for extended comments and support.

This project required enormous amount of data work. I would like to thank David Brown and John Earle for their academic guidance when I was working as their research assistant. Thanks to you I learned a lot about the Ukrainian data and acquired first hand-on experience in academic work. I am indeed grateful to Vladimir Mamontov who introduced me into the hectic world of Ukrainian enterprise statistics. Your comments and insider advice were really invaluable! I would also like to thank Valentina Smal for her help with the geographic part of the data.

I would like to acknowledge my colleagues at Kyiv School of Economics (especially Denis Nizalov and Olexander Shepotilo) for your willingness to share your expertise during the seminars and countless discussions, and also the KSE stuff for providing me with an office space, and administrative and financial support.

This work would never been done without people outside of academia who are very important to me. First of all, I feel indebted to my wife, my colleague and the best friend Hanna for her constant encouragement, spousal support and love that kept me floating while away from home. Also, I will always keep in my memory many letters from my sons, Bohdan and Anton, who not only patiently waited for my return, but also learned how to read, write, and type on a computer while I was far from home. You brought a lot of happiness to my life. I could never have gone this far if did not think of you, kids!

My parents and my wife's parents deserve a special gratitude for their relentless support and inspiration. I appreciate your understanding and patience when I was too busy and did not return phone calls and emails. Mom, I finally did it!

I really appreciate help and emotional support of my friends in Lexington, Kathryn and Jeremy Cox, Jaimy and Emily Tillett, and Jack and Kay Frost. Life was so much easier with you being around, guys!

Last, but not least, I would like to thank Terry Pratchett. When I had a free moment to read, your "Diskworld" series of books kept me sane.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

Economic activity is spatially concentrated only in a number of locations, rather than evenly distributed all over the place. This fact has been established by studies in economic geography, regional and urban economics and history. Fujita and Thisse (2002) provided numerous examples of localization of the economic growth in different countries. It is evident that the phenomenon of extreme variation in the land use is not pertinent to a particular country or a historical epoch, but rather is broadly observed throughout the history of the human civilization.

In this work I will study industrial agglomeration in Ukraine using a rich panel dataset for the period of 2001 to 2005. I explore whether the spatial distribution of economic activity in the formerly planned economy follows the same principles as in more developed countries. I will show that major findings from the Western literature about agglomeration effects are supported in the case of Ukraine, with an important extension towards the effects of the ownership structure. I will also study how aggregation of industrial and territorial data may affect the results.

Firms in some industries not only tend to spatially concentrate, but also chose locations close to other firms in the same industry forming industrial clusters. Clusters in such resource-oriented industries as coal mining, gas and oil extraction, quarrying, or metal welding, are the most evident. Nonetheless other manufacturing industries in different countries reveal significant geographic concentration as well. Examples include concentration of car makers around Detroit area in the beginning of the twentieth century, watch-making districts in Switzerland, or apparel producers in Northern Italy. Following Marshall (1890), these effects are known as Marshal scale economies. Since the effects are external to the firms but internal to the firm's industry and location, they are also known as localization economies. Marshall (1890) noticed that location of firms of the same industry in the same place generates additional effects for their mutual productivity. He explained benefits of collocating by an easier access to the common labor pool and specialized input services, an access to the same market for many producers, and sharing the ideas between firms and employees leading to information spillovers and accumulation of industry-specific human capital. Fujita and Thisse (2002) also noticed "a lock-in effect" of initial location decisions, when an industry, once placed in a certain location, tends to attract new plants that "want to congregate to benefit from a larger diversity of activities and a higher specialization." Henderson (1974) formalized these observations into a general equilibrium model which predicts that cities will tend to specialize in production and trade of a certain good (industry) or a bundle of goods and trade, whereas providing non-tradable services on the spot. The model also predicted that cities may vary in size depending on the level of scale economies in production of that particular good. Clustering of all firms form the same industry in just one place is a theoretical extreme, but empirical studies support both the existence of Marshall external scale economies in various industries and tendency of some industries to a greater geographic concentration compared to others.

The literature also distinguishes another type of external scale effects. These effects are external not only to the firm, but also to the industry, though internal to the firm's location. Following Jacobs' (1969) work, such effects are dubbed as urbanization economies. Jacobs considered a city as a congregation of diverse firms that are actively engaged in mutual work, constant innovating and invention of new activities, which gradually replaced outdated ones. Ability of a city to maintain these Schumpeterian churning processes, provide variability of services and industrial diversity, and support the exchange of ideas, in Jacobs' view, were necessary conditions for a sustainable urban growth.

The distinction between localization and urbanization economies is important from the policy point of view. When a firm expects to benefit from the localization economies, it will tend to locate in a city where other firms in the same industry are already present. As Duranton and Puga (2001) show in their model of "nursery cities," such behavior is more typical for industries with established production process and standardized technology. Specialization of a city in a selected technology (or a set of closely related technologies) lets firms in that city to save on production costs and fully take advantage of scale externalities, as the city is able to provide firms with specialized labor force and enhance knowledge spillovers between firms. As a result, a city which hosts such industries may choose to follow policies designed to attract other firms in the same industry from other regions. On the other hand, firms may also seek for benefits from urbanization economies and to locate in a city with a more diversified industrial profile. According to Duranton and Puga (2001), this is more common for firms in young industries at an early stage of their development, ones that are still "seeking for the best technological process." Production costs for firms engaged in such "experimental activities" are probably higher since urban diversity tags its price in terms of amenities costs and congestion. On the other hand, diversity provides "cross-fertilization" of ideas and production methods, which are essential for young firms. Fujita and Thisse (2002) mentioned that "observed spatial configuration of economic activities is the result of a complicated balance of two opposing forces, that is, agglomeration (or centripetal) forces [that push consumers and firms together], and dispersion (or centrifugal) forces [that pull them apart]." Hence, there are two possible avenues for urban development. If the area is dominated by well-established industries, urban planners may choose to attract firms from the same industries to benefit from localization economies within those industries. On the other hand, urban planners may want to support industrial diversity in the area and thus provide more grounds for urbanization effects. These two goals are not necessarily mutually exclusive. Other studies (for example, Black and Henderson (2003)) indicate that cities specializing in only one or a few industries tend to be smaller, on average. Henderson (1974) also mentioned that the city size depends on the ability of the industry the city is specialized in, to benefit from localization effects. Therefore cities aimed at fast and sustainable growth, are expected to attract not only more firms from within the same industry, but also more various industries. It is important for urban planners to understand what type of industries prevail in the area to make decisions about further investment goals and city development as a whole.

Most large firms which formed the core of the Ukrainian economy were established decades ago. Decisions about investment and location of firms in the Soviet planned economy were often based on different principles than the market economy would imply. David Dyker (1983) names "tendency to overbid for investment resources," "a lack of coordination between state bodies responsible for investment decisions," and "elements of operational inefficiency at the design and construction stages" among others as factors that have ultimately determined the composition and location of production assets in the former Soviet Union. Dyker points out that obsession with constructing huge plants and the call for development of unpopulated areas were declared in the doctrinal principles of the Soviet economy. As a result, manufacturing industries were concentrated and often located in single company-towns, with big plants dominating local labor markets. In 1971, about ten per cent of industrial firms employed about sixty percent of the total industrial personnel and produced about two thirds of the total industrial output in the USSR, whereas the average size of a manufacturing enterprise exceeded 1000 employees. (Dunaev (n.d.)). The entire system of production in the Soviet Union was determined by the central planners, and firms' discretion in the choice of the most appropriate suppliers and customers was limited. Dyker (1983) provides several examples (case studies) where firms in different republics located thousands of miles apart participated in the same production process.

I am not implying that external scale economies were absent in the Soviet Union altogether. Not only in the established urban areas, but even in green field developments, the public infrastructure and service industries soon followed the leading firms thereby reducing production costs. Organization of firms from several supplementary industries into "territorial-production complexes" (Lonsdale (1965)) was also designed to minimize the costs. Hence, it is possible to assert that both urbanization and localization economies were present in the Soviet economy, at least to some extent. Nevertheless, a distinctive feature of the Soviet economic system was the absence of markets for land, goods and services, and, as a result, factor and goods prices were distorted. Firms were bound by pre-determined production targets (as stated in the Central Plan) and behaved as cost-minimizers given their set of capital assets and struggling for their share of supplies and raw materials. On the other hand, the potential market area for a Soviet firm could be the entire Soviet Union, as the central planners had complete discretion over output flows. Since transportation costs were not a bounding factor for planners, some firms and even entire industries were heavily subsidized by the state, especially when inter-industry linkages involved long-haul shipment of materials and ready products across the country. Such implicit subsidies do not permit correct estimation of agglomeration effects in command economies.

The transition period in the USSR and countries of Eastern Europe began in late 1980s. Deterioration of the socialist economic system, beginning in mid-eighties and reaching its culmination with the collapse of the Soviet Union in 1991, gave a rapid start to mass privatization in new independent states. The Ukrainian data show that the first privatization deals in Ukraine were dated in early 1992. At that time the process was going at a slower pace compared to the neighboring countries of Eastern Europe and Russia. Paskhaver et. al (2003) assert that the first stage, "mass privatization," lasted from 1994 to 1998. During that period, mostly small firms were privatized through share distribution among managers and workers. Brown et. al (2006) claim that such privatization scheme resulted in a low concentration of ownership and substantial levels of state control. By the end of the nineties, the next stage of privatization began, when big shares in large manufacturing firms were traded at auctions and stock exchanges or through direct sales. Paskhaver et. al (2003) reports that by the turn of the century private firms employed more than a half of the labor force and produced almost two-thirds of the total output in Ukraine. At the same time, privatization has slowed down its pace. Brown and Earle (2007) report that in a sample of manufacturing firms which were state-owned in 1989, more than seventy percent have been privatized by the end of 2000, with a steady increase of the share of majorly foreign firms from zero to 2.5% in 2005. The focus has shifted towards the transfer of the remaining shares of the formerly state firms and privatization of remaining large firms. Today the process is almost complete. According to Derzhkomstat (2006), only two to three per cent of all firms in manufacturing were completely state-owned at the beginning of 2006.

I would like to mention two important distinctions of transition economies from more developed countries. First, the share of state owned firms is significantly greater at the early stage of transition. Second, all countries evidenced a rapid and massive transfer of property rights from the state to private economic agents. Aslund (2000) claimed that the speed and scope of privatization in the countries of the former Soviet Bloc, including Ukraine, were unprecedented: over a short time span of ten to fifteen years, more than two thirds of enterprises in every transition economy completely changed their ownership type. On the other hand, this time was probably not enough for all firms to quickly adjust to such changes. Privatization and price liberalization in the nineties initiated emergence of markets for labor and resources (the land market has yet to be liberalized). This was a period when manufacturing firms started to face and act in the true market environment, and the role of the state in production decisions was finally eradicated. The previous distortions in prices were soon revealed. It became evident that previous linkages with other firms were not always feasible any longer due to increased transportation costs, customs regulations (if they were located in different countries), or the changes in the output mix of the former suppliers. Hence, the firms had to change their production plans or to reallocate to a different sector. According to Derzhkomstat (2006), the entire decade of nineties in Ukraine was marked by the output drop; the national output started to increase in 2000 for the first time since Independence. Very often large firms designed as a part of the all-Soviet production chain had to downsize, since the Ukrainian market was too slim for them. This process was exacerbated by a drop in population and, therefore, in the total labor force. That is why one of the effects of the privatization was a decrease in the firm size: large firms split into smaller entities, and newly established firms were smaller than the former Soviet plants. Very rapid changes in the business environment, downscaling of the former Soviet enterprises and reallocation of production were the most pronounced features of the first phase of transition in Ukraine.

Economic theory states that private firms are usually more responsive to rapid changes in the business environment. The data I use let me to determine the ownership type for each enterprise in every year. I believe it is important to control for the ownership in agglomeration studies in transition economies, because the speed of adjustment and the degree of involvement in relationships between firms may vary between private and public firms. Even though privatized firms are bound with production assets in a certain location, they are still expected to have more discretion in the choice of the output mix and business partners. Newly created firms have even greater degree of freedom in this respect. That is why one of the underlying hypotheses in my study is a claim that private firms benefit more from agglomeration economies. As I show in my work, this statement appears to be true: private firms seem to benefit more from agglomeration economies, and foreign private firms enjoy even greater benefits than domestic private firms.

My work generally supports findings from other agglomeration studies based on data from various industries in the developed countries (see Rosenthal and Strange (2004) for an extended review). I show that agglomeration economies are present in Ukrainian manufacturing industries, and the scope of the effects seems to be similar to those in other studies. However, the direct comparison of findings across studies is often not possible due to different scale of agglomeration measures and the level of data aggregation. Even though localization or urbanization economies are external to an individual firm, the ultimate goal is to explore how they affect the firm's productivity. For earlier studies, firm-level data were not available.[1] Using the terminology of Ellison and Glaser (1997), the data were usually aggregated into "an industry-region cell", where "the industry" was taken at two-digit code of a national classification, and "the region" was a state, a prefecture, a metropolitan area or a city. The size of the cell in terms of employment or number of firms was large, which raised concerns in the literature about the agglomeration effects measured at this level having the same effect for all firms.

A standard proxy for urbanization economy is the size of a city in terms of total population, total employment or total manufacturing employment. The localization economy is measured as the size of an industry-region cell, in terms of employment or total number of plants within the same industry and geographic area. Even though such variables were widely used, there was no consent in the literature if they are most appropriate. The original idea of Jacobs (1969) implied that urbanization effects result not from a city size alone, but from diversity within a city. Carlino (1979) mentioned that population scale was not the most appropriate measure since it could incorporate both agglomeration economies and diseconomies. Sveikauskas (1975) was concerned that aggregating the data within such large industry-area cells may conceal many relationships taking place locally, such as division of labor specialization between firms. Moomaw (1983) have shown that population scale is an appropriate measure once the estimation is controlled for other important factors, such as public infrastructure, population density, or industry. He warned against aggregating all industries into the same equation and advocated separate estimations. Nakamura (1985) originally measured the both localization and urbanization effects at the level of two-digit industries in Japanese prefectures. At the same work he re-estimated his model at the level of cities and found similar results, leaving aside an issue of agglomeration effects for the firms located outside of urban areas. Moomaw (1998) compared estimates for two- and three-digit

[1] Refer to Eberts and McMillan (1999) for an extended review of earlier agglomeration literature.

industry levels in MSAs and also found general similarities in the results for most industries. He also warned against excessive disaggregation since in that case, it is possible that not all regions will have all industries, which may result in the loss of degrees of freedom. Chiccone and Hall (1996) have shown that despite the similarities in the geographic size of an industry-region cell, the spatial density of activity within the cell seems to be an important factor for productivity. Finally, Ellison and Glaser (1997) have theoretically established an "observational equivalence between the effects of natural advantages and spillovers [between firms]." In their later work (Ellison and Glaser (1999)) they estimated that "at least half of the concentration is due to natural advantages," whereas in a number of geographically concentrated industries individual interfirm spillovers matter, but cannot be measured with aggregated data. To conclude, data availability in earlier studies precluded their authors from running agglomeration analysis at more local levels or for individual firms. Nevertheless, they acknowledged that agglomeration economies are determined not only by location of firms within industry-region cells, but also by more local factors, down to relationships between individual firms. These factors may act both pro- and against agglomeration, but their exposure can hardly be studied with too aggregated data.

Aggregation of data into industry-region cells leads to an implicit assumption that all firms within the cell share similar characteristics and respond similarly to the same external effect. This assumption seems to be too strong. Even though all firms within an industry-region cell face the same localization and urbanization effects, they may respond differently. Such differences were stressed upon in a number of more recent studies. In their extended review of this literature, Rosenthal and Strange (2004) start with a benchmark model where agglomeration is presented as an external shift factor, which in turn depends on interactions between individual firms in the geographic, industrial and temporal space. The authors admit that data requirements to estimate such a model are enormous and probably infeasible at the moment. At the same time, increasing availability of firm-level data allows making inference about effects of agglomeration economies on individual firms.

The first direction of the research was related to availability of geo-referenced firm-level data which allow measuring agglomeration effects based on exact distances between firms. Proponents of this approach claim that it lets them to alleviate a so-called Modifiable Areal Unit Problem (MAUP) problem, when aggregating the data into larger geographic units may lead to misspecification issues and spurious spatial correlation between the units. Several such studies let us conclude that agglomeration effects tend to attenuate much sooner than it was believed before. If any spillovers between firms take place, in most cases such spillovers are very local in nature and usually do not stretch far beyond city limits. Rosenthal and Strange (2003) explored location decisions of new US firms contingent on the geographic concentration of their own industry and found that agglomeration effects start to attenuate already after five to ten miles. Duranton and Overman (2005) constructed a localization measure based on exact inter-firm distances within the universe of UK manufacturing firms and found that “localization of three-digit industries is important at small scales (between zero and fifty kilometers) and at a more regional level (between eighty and 140 kilometers).” Finally, Cainelli and Lupi (2008) also measured exact distances between Italian firms and made similar conjectures about decreasing probabilities of agglomeration effects with the distance within Bayesian framework. All these studies indicate that physical distances between firms matter a lot for localization economies. What is more important, the threshold distances do not necessarily correspond to the size of political or statistical regions which usually determine the level of data aggregation.

Another component in the model of Rosenthal and Strange (2004) was the “industrial distance”, tentatively defined as the frequency of relationships between firms from different industries. Authors conjectured that *ceteris paribus*, firms in similar or complementary industries tend to locate closer to each other. A standard way of dealing with similarity of industries is using a national classification of industries, such as NAICS in North American studies or NACE in European works. For example, Moomaw (1998) compares agglomeration effects within three-digit industries with the corresponding two-digit aggregate. A problem with national classifications is that they define industries based on similarities in their outputs rather than production processes and ignore supply-chain relationships between smaller industry groups within a larger aggregate. As a result, a two-digit industry may combine quite diverse three-digit components, and the agglomeration at higher aggregation levels may be very different from the agglomeration within industries at lower levels. However, most of the time statistics are only available for large industrial groups, and researchers have little discretion about this measure.

The firm-level data are collected by national statistics offices (as in Henderson (1986, 2003) and Brown et al. (2006)) or independent institutions (as "Dun and Bradstreet Marketplace Database" used in Rosenthal and Strange (2003)). The data sets usually include statistics on output, major inputs, such as labor, capital assets and materials measures, total wage bill, as well as firm information sheets with detailed industry codes, territory codes, ownership codes, etc. Due to non-disclosure issues, firms' names and addresses are usually omitted or restricted, which makes it difficult to geocode each observation. On the other hand, industry and territory codes are much more detailed compared to the earlier studies. For example, the industry codes are available up to four or even five digits of the national industry classification, and the territory codes are disaggregated down to the level of a zip-code, county or a smaller labor statistical area. Even though a researcher still "transforms dots on a map into units in boxes", as Duranton and Overman (2005) rightfully mentioned, such industry-region cells are much smaller. Comparing the estimates of localization effects measured at various levels of aggregation within the industrial and geographic space — varying a cell size — may help understand the effects of aggregating the data and the scope of agglomeration.

In this work I measure the agglomeration effects using Ukrainian plant-level data from manufacturing for the years 2001 through 2005 and compare the effects at various degrees of aggregation in the geographic and industrial space. Following the logic of geo-coded studies (Rosenthal and Strange (2003), Duranton and Overman (2005)), I first measured agglomeration at the smallest geographic and industrial scale given the data restrictions. Then I compared these estimates with those which measured the effect within larger industry-region cells. My findings are two-fold. When agglomeration is measured at the same level of industrial aggregation but within a larger geographic area, the effect is similar. This result supports a hypothesis of the spatial attenuation of the effects. However, when I use the same level of territory aggregation, but expand the industry definition from a three-digit to a two-digit industry cell, the effects vary by sectors, indicating different relationships within sectors and different response to agglomeration effects compared to smaller industry groups.

The rest of the work develops as follows. In the next chapter I describe the data. In the third chapter I provide descriptive statistics and analyze the patterns in agglomeration economies and spatial concentration in manufacturing sectors. I also concentrate on the description of the machinery and the higher technology industry groups. In the fourth chapter I focus on the analysis of agglomeration economies in these two industries, as well as on the relationship between agglomeration economies and firms' ownership strucutre. In the fifth chapter I explore the issues of data aggregation and attenuation of agglomeration economies with the distance. The sixth chapter concludes.

Chapter 2. Data description

2.1. Production Function Controls

The data in this work are structured similarly to the U.S. Longitudinal Research Database (LRD) described in Henderson (2003) and Dun and Bradstreet's Marketplace file (D&B) used in Rosenthal and Strange (2003). The data used here are establishment level, longitudinal and collected annually. The data span the years 2001 to 2005. The data are restricted and not available for general public use. [2]

The particular choice of the period was determined by the following factors. First, 2001 was the year when the first stage of major structural changes in the Ukrainian economy came to an end. The majority of enterprises were privately owned, relative financial stability has been achieved, and the national product and manufacturing output started to grow. According to OECD (2007), real GDP growth in Ukraine averaged 7.3% over 2000-2006, in contrast to the severe recession during early transition period in the nineties. Second, beginning 2001 the data became internally consistent over time. Even though some variables were available for 1998-2000, that period was marked by major changes in the national industrial classification system and a reform of accounting standards; hence the definition of key variables was different before 2001. Finally, data for later years were not available at the time when I composed the dataset.

The data include both output and input measures such as net sales, total assets, employment, and the wage bill, and derive from annual the Enterprise Performance, the Financial Results, and the Balance Sheet statements that all firms must submit to the National Statistics Office (Derzhkomstat).

Output measures come from the Financial Results Statement. The data provide several possible definitions of output (sales, sales adjusted for materials, gross vs. net sales). Small firms submit a simplified version of the statement. Gross sales (before taxes) and material costs were not available for the small firms and branches of the big firms. I experimented with imputing the missing data based on output to labor ratio, but the resulting coefficients for all production factors turned out to be out of reasonable

[2] I would like to appreciate the support of Kyiv School of Economics in the data support.

range. Since net sales have shown adequate performance in the preferred specification I decided to use this variable.

Employment was taken from the Enterprise Performance Statement. Employment is calculated as the “year-average number of enlisted employees”. I believe that this measure is superior to the nominal head count and comparable to the number of hours worked as used in Henderson (2003). For every firm, there are data for staffed and non-staffed employees. The non-staffed employees are considered those who worked part-time or as independent contractors during the year. The Enterprise Performance Statement is the only one that also provides additional information about a firm’s branches and affiliates whenever they are present. However, only employment and net sales are available for each branch, and employment is not divided into staffed and non-staffed workers for branches. Hence I decided to use the total employment both for firms with branches (for the establishments) and for stand-alone firms.

The annual Balance Sheet Statement was the source for the capital variable. The capital measure is based on the nominal end of year value of the tangible assets. To overcome the difficulty with the lack of the capital measure for branches, I decided to use intrafirm capital – labor ratio as a basis for imputing capital levels for each establishment within a multiplant company. I calculated the capital – labor ratio for the entire firm and multiplied it by the value of employment in every firm’s branch. The total number of firms with branches in manufacturing increased from 442 (0.96%) in 2001 to 668 (1.28%) in 2005. The branches constituted between three percents of all establishments in 2001 and six percents in 2005, whereas the total share of imputed capital in branches increased from twenty five percents in 2001 to thirty seven percents in 2005. Henderson (2003) had an access to establishment-level measures of capital and expressed explicit concern about using imputed data. It is also possible that capital-labor ratios will be different for “managerial offices” and “production units” as well as for firms located in rural areas vs. urban locations. I understand grounds for such concerns, but I could not get branches’ capital otherwise.

Every annual Enterprise Performance Statement is accompanied by a comprehensive description file. The file contains data on the establishment ID, mother firm ID (if applicable), property type, organizational form type, industry code, and territory code. Based on this file I was able to link each establishment with the corresponding codes in the industry and territory classifications. I also created a separate indicator variable “subsidiary” which is one if an entity (a branch or a separate firm) shows another firm as its “mother company” or a “head office”. The structure of management in Ukrainian companies is rather complex. I believe that various vertical relationships can be captured by this variable.

2.2. Industrial Classification

The classification of industries in Ukraine (KVED) was introduced in 1998-2000 and coincides with European industrial classification NACE rev.1 and International Standard Industry Classification (ISIC) at the level of four digits. For this work, I have chosen the entire set of manufacturing industries. The manufacturing sector is coded with the letter “D.” The sector is divided into 14 subsectors, ranging from “DA” to “DN”. Each subsector combines one or two two-digit divisions; each is further divided into three-digit groups and four-digit classes. The assignment of firms to a particular industry code is based on the particular distribution of shares of the firm’s output among sectors. The entire list and the hierarchical structure of subsectors, two- and three-digit industries are shown in Table 2.1.

For every firm in the sample, I have obtained its three-digit and two-digit numeric codes and a two-letter subsector code. Whenever I encountered missing values I tried to fill them in with values from the previous and following periods if they were the same. I had to drop thirty six branch observations in 2003-2005 due to the lack of the industry code at the level of at least three digits.

For the study of agglomeration economies, I have chosen machinery and high-tech industries aggregated at the three digit level. This particular choice of industries and the aggregation are influenced by compatibility with available studies, such as Henderson (2003). Another rationale is that machinery represents an established industry whereas high-tech is a new and growing. Such division is in line with Duranton and Puga’s (2001) theoretical model. The industries in the machinery group are KVED 291 (Manufacture of machinery for the production and use of mechanical power, except aircraft, vehicle and cycle engines), 292 (Manufacture of other general purpose machinery), 294 (Manufacture of machine tools), and 295 (Manufacture of other special purpose machinery). The high-tech industries are KVED 296 (Manufacture of weapons and ammunition), 300 (Manufacture of office machinery and computers), 321 (Manufacture of electronic valves and tubes and other electronic components), 331 (Manufacture of medical and surgical equipment and orthopedic appliances), and 353 (Manufacture of aircraft and spacecraft). The first three-digit industry in the list, "Manufacture of weapons and ammunition" (KVED 296) was excluded from the similar study in Hederson (2003). In my sample, this industry includes only 112 observations (between fifteen and twenty six observations per year), and inclusion or exclusion of it does not significantly change the result, which is why I decided to keep it in the sample.

For the study of attenuation effects, I used all two-digit subsectors. Two subsectors, DC ("Manufacture of leather and leather products") and DF ("Manufacture of coke, refined petroleum products and nuclear fuel") have a very small number of observations (1.5% and 0.5% of the total sample, respectively). Hence, I decided to merge them into the subsectors DB ("Manufacture of textiles and textile products") and DG (""Manufacture of chemicals, chemical products and man-made fibers), respectively.

The issue of comparability of industry classifications between countries remains open. The European industry classification (NACE) and, consequently, the Ukrainian industry classification (KVED) are not directly comparable with the North American classification (either SIC or NAISC), nor with the Japanese classification. Nevertheless, two particular industry groups (machinery and high-tech) are similar at the level of four-digit codes and detailed description. The similarity of definitions of these two groups provides necessary grounds for comparing my results with those of Henderson (2003). Also, the general principles of classification are preserved internationally, and I am still able to make comparisons for the majority of subsectors and two-digit industries.

2.3. Territory classification

The territory system in Ukraine reflects complex rules of administrative subordination. The entire country is split into twenty four oblasts (which are comparable to US states), one autonomous republic of Crimea and two cities with the special status – Kyiv, which is the national capital (formerly Kiev), and Sevastopol in Crimea. These units constitute the first, upper level of the administrative division of the country. Each oblast is further split into a number of rural raions (similar to US counties) which are administrated from the oblast capital city. Big raion capitals and sometimes other large cities within raions are also administrated from the oblast capital city and have a status of cities with oblast administrative subordination. This status is assigned based on the city population (which should be at least 50,000 people), and the level of the industrial and cultural development. The number of such cities ranges from one per oblast in the predominantly rural Western part of Ukraine to twenty eight in heavily industrialized Donetsk oblast in the East. Rural raions and cities of oblast administrative subordination constitute the second, medium level of the administrative division. I perform my analysis at this level. There are 490 rural raions, 177 cities of the oblast subordination and two cities of the national subordination, which totals to 669 regions. Whenever I refer to these regional units I call them all “raions.”

Every establishment in the database has a territory code according to the National system of administrative units (KOATUU). Based on this code I created the variable “Raion ID”. Only three observations were dropped due to the lack of the territory code. I also created an additional variable “Urban” to indicate establishments located in cities of oblast subordination, Kyiv and Sevastopol.

To increase compatibility with Western studies, especially Henderson’s (2003) findings, I created an analog of Metropolitan Statistical Areas for Ukraine. Since commuting patterns are not readily available for Ukrainians raions, I used a simplified approach. I took neighboring raions within sixty kilometers from cities with the population over fifty thousand people, in a hierarchical order. If a raion is within sixty kilometers of multiple cities, it is “assigned” to the city with the greatest population. The population counts are taken from 2001 Population Census, and the matrix of distances is based on air distances between the administrative centers of each area.[3] I believe that this approach is a good measure of MSA’s in Ukraine given data availability at the moment. According to Ukrainian administrative laws, a population threshold of fifty thousand is considered the smallest boundary to call a municipality a city. From private conversation with representatives of Ukrainian Association of Employers, sixty kilometers is deduced as the maximum transportation distance for working commuters. As a result of such exercise, I ended up with fifty six “quasi-metropolitan” (QMSA) areas for Ukraine

[3] The matrix of distances was kindly provided by Regional Studies Group at Kyiv School of Economics

covering 537 raions and cities (about eighty percent of the total amount of administrative areas). It is remarkable that about ninety five percent of firms from my database are located in one of these QMSA's. See Figure 2.1 for the map of QMSA areas in Ukraine. I also experimented with fifty kilometers transportation distance, which resulted in 465 areas covered by sixty one QMSA's, but it did not significantly affect my results. For the analysis of agglomeration economies and attenuation of agglomeration effects I used only establishments located in QMSA's.

2.4. Ownership variables

For any study of transition economy, the establishment's ownership type is an important factor. I used data from State Property Fund database and Balance Sheet Statements to construct a panel of two control indicator variables, "primarily domestic owned" and "primarily foreign owned." Brown et.al (2006) used this pair of variables to examine the productivity effects of privatization. The State Property Fund data were used to construct annual shares of private vs. public ownership for each firm since 1992. The dataset included both firm records and records about stand-alone public objects subject to privatization, such as unfinished construction sites or parts of state-owned buildings. Luckily, SPF data also contained the value of the statutory fund for each privatized object. I matched statutory fund variable from the SPF database with the same variable in the database of the State Registry and Balance Sheet Statements, and decided to denote an object "a true firm" if the discrepancy in the value of the statutory fund between different sources did not exceed 10%. The resulting table of annual private shares was amended with a number of firms banned from privatization by the Government orders.[4] Such firms were marked as state-owned in a given year. If the private share if the private share in a given year was below one half such firms were also marked as state-owned. All other firms were marked as majority private. Then I added foreign ownership shares from FDI dataset provided by Derzhkomstat and marked all majority private firms as primarily domestic owned or primarily foreign owned depending on which share was higher in each year.

---

[4] I would like to thank David Brown and his Research Assistant, Yevgeniya Shevtsova, for providing the data.

Thus, each observation has establishment ID, subsidiary indicator, three-digit industry KVED code, raion ID code, QMSA ID, output, value of total assets, total employment, urban or rural indicator, and a pair of ownership indicators. Refer to the Table 2.2 for the sample composition details. Table 2.3 shows the composition of the manufacturing two-letters subsectors, whereas the composition of the sample of machinery and high-tech industry groups is presented in Table 2.4.

2.5. Descriptive Statistics

I present the descriptive statistics of the major variables using the pooled sample in Table 2.5. The standard deviations are very large, which may suggest substantial heterogeneity in the data. Therefore in all other aspects I present the descriptive statistics by sectors and, when variation between periods is substantial, by years. As I also mentioned earlier, Ukrainian firms submit one of two different types of Balance Sheet and Financial Result statements depending on the firm size. A firm is considered big if it employs over fifty people. In Table 2.6, I present the shares of large firms by subsectors. On the average, eighteen percent of establishments are considered large, ranging from seven percent in DD ("Wood and wood products") and DE ("Pulp, paper, and publishing") to over thirty percent in DM ("Transport equipment"). At the same time, on average, large firms employ eighty-six percent of total manufacturing employment, possess ninety three percent of total assets, and produce eighty-nine percent of total output in the sector (refer to Table 2.7 for the distribution of the employment, capital and output by subsectors of manufacturing). I present the means for the major variables (employment, capital, output and urban indicator) large manufacturing firms in Table 2.8, and for small firms in Table 2.9.

In Table 2.10 I show the same descriptive statistics for machinery and high tech industry groups. It seems that the average size of small firms does not significantly change, whereas large firms appear to shrink over the observation period. This phenomenon is especially pronounced for high-tech firms, which display a fifty percent downsizing over the five years. In addition to the major estimation variables I also present the means of the ownership indicators. It appears that significantly fewer firms in high-tech are private compared to machinery and economy in general. An average majority state owned firm is also larger than an average majority private firm, as indicated in Table 2.11. This fact may be explained by a greater share of "strategically important" large firms which specialize predominantly in high tech industries, and privatization of which is banned or restricted by the state. Only one to four percent of private firms in both industry groups are majority foreign owned, as it is evident from Table 2.10. The table also demonstrates that foreign owners seem to invest mainly in large firms. In Table 2.12 I present the average establishment employment for majority foreign owned and majority domestic private firms. Even though majority foreign firms seem to be greater, on average, this is especially true for firms in machinery rather than in the high tech industry group. Starting in 2002, establishments in foreign-controlled machinery firms are even greater in size than majority state-owned firms. This fact may indicate that the share of large private firms tends to increase over time.

Firms located in urban and suburban raions may differ in size and production factor use. This difference may affect the precision of imputed capital values for firms' branches. In Table 2.13 I present the average capital-labor ratios for establishments located in urban and suburban raions. In general, the K/L ratios tend to grow in time, since the nominal value of total assets tends to increase, whereas employment has fallen over the estimation period in nearly all subsectors. At the same time, for most subsectors except DG ("Chemicals, man-made fibers") and DI ("Other non-metallic mineral products), the ratios are greater for suburban firms compared to urban ones. In one more sector, DE ("Pulp, paper and publishing), the ratios are greater in four years out of five, but the difference is not that pronounced compared to the former two. On the other hand, urban establishments are generally larger than suburban ones, as I show by the size of average employment and average total assets in Table 2.14 and Table 2.15, respectively. Further analysis shows that it is not always possible to correct for difference in capital-labor ratios between urban and rural establishments when imputing the capital values for branches, because for 1,502 manufacturing branch observations both the main firm sector and the main firm urban status differ from those of the branch. In other words, the main firm reports a different urban or rural location and a different industry sector than its branches. Nevertheless such branches constitute less than one percent of all observations, employ less than three percent of the entire employment and use less than five percent of the entire capital, which let us hope that the imputation error is not significantly large.

Finally I present the descriptive statistics for the largest cities. Firms located in cities of Kyiv, Donetsk, Dnipropetrivsk, Mariupol, Zaporizhia, and Kharkiv employ twenty eight percent of the total manufacturing employment in 2001, and this share increases to thirty one percent by the end of 2005. I have chosen these six cities because each of them employed over three percent of the labor force alone. The descriptive statistics is shown in Table 2.17 through Table 2.21.

Tables

Table 2.1, Subsectors, Divisions, Groups and Classes of Sector “D” (Manufacturing)

PSEK	KVED2	KVED3	Title
DA	15	15.1	Production, processing and preserving of meat and meat products
		15.2	Processing and preserving of fish and fish products
		15.3	Processing and preserving of fruit and vegetables
		15.4	Manufacture of vegetable and animal oils and fats
		15.5	Manufacture of dairy products
		15.6	Manufacture of grain mill products, starches and starch products
		15.7	Manufacture of prepared animal feeds
		15.8	Manufacture of other food products
		15.9	Manufacture of beverages
	16	16.0	Manufacture of tobacco products
DB	17	17.1	Preparation and spinning of textile fibres
		17.2	Textile weaving
		17.3	Finishing of textiles
		17.4	Manufacture of made-up textile articles, except apparel
		17.5	Manufacture of other textiles
		17.6	Manufacture of knitted and crocheted fabrics
		17.7	Manufacture of knitted and crocheted articles
	18	18.1	Manufacture of leather clothes
		18.2	Manufacture of other wearing apparel and accessories
		18.3	Dressing and dyeing of fur; manufacture of articles of fur
DC	19	19.1	Tanning and dressing of leather
		19.2	Manufacture of luggage, handbags and the like, saddlery and harness
		19.3	Manufacture of footwear
DD	20	20.1	Sawmilling and planing of wood; impregnation of wood
		20.2	Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board, fibre board and other panels and boards
		20.3	Manufacture of builders carpentry and joinery
		20.4	Manufacture of wooden containers
		20.5	Manufacture of other products of wood; manufacture of articles of cork, straw and plaiting materials

Table 2.1 (continued); Subsectors, Divisions, Groups and Classes of Sector "D"

PSEK	KVED2	KVED3	Title
DE	21	21.1	Manufacture of pulp, paper and paperboard
		21.2	Manufacture of articles of paper and paperboard
	22	22.1	Publishing
		22.2	Printing and service activities related to printing
		22.3	Reproduction of recorded media
DF	23	23.1	Manufacture of coke oven products
		23.2	Manufacture of refined petroleum products
		23.3	Processing of nuclear fuel
DG	24	24.1	Manufacture of basic chemicals
		24.2	Manufacture of pesticides and other agro-chemical products
		24.3	Manufacture of paints, varnishes and similar coatings, printing ink and mastics
		24.4	Manufacture of pharmaceuticals, medicinal chemicals and botanical products
		24.5	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations
		24.6	Manufacture of other chemical products
		24.7	Manufacture of man-made fibres
DH	25	25.1	Manufacture of rubber products
		25.2	Manufacture of plastic products
DI	26	26.1	Manufacture of glass and glass products
		26.2	Manufacture of non-refractory ceramic goods other than for construction purposes; manufacture of refractory ceramic products
		26.3	Manufacture of ceramic tiles and flags
		26.4	Manufacture of bricks, tiles and construction products, in baked clay
		26.5	Manufacture of cement, lime and plaster
		26.6	Manufacture of articles of concrete, plaster and cement
		26.7	Cutting, shaping and finishing of ornamental and building stone
		26.8	Manufacture of other non-metallic mineral products

Table 2.1 (continued); Subsectors, Divisions, Groups and Classes of Sector "D"

PSEK	KVED2	KVED3	Title
DJ	27	27.1	Manufacture of basic iron and steel and of ferro-alloys
		27.2	Manufacture of tubes
		27.3	Other first processing of iron and steel
		27.4	Manufacture of basic precious and non-ferrous metals
		27.5	Casting of metals
	28	28.1	Manufacture of structural metal products
		28.2	Manufacture of tanks, reservoirs and containers of metal; manufacture of central heating radiators and boilers
		28.3	Manufacture of steam generators, except central heating hot water boilers
		28.4	Forging, pressing, stamping and roll forming of metal; powder metallurgy
		28.5	Treatment and coating of metals; general mechanical engineering
		28.6	Manufacture of cutlery, tools and general hardware
		28.7	Manufacture of other fabricated metal products
DK	29	29.1	Manufacture of machinery for the production and use of mechanical power, except aircraft, vehicle and cycle engines
		29.2	Manufacture of other general purpose machinery
		29.3	Manufacture of agricultural and forestry machinery
		29.4	Manufacture of machinetools
		29.5	Manufacture of other special purpose machinery
		29.6	Manufacture of weapons and ammunition
		29.7	Manufacture of domestic appliances n.e.c.
DL	30	30.0	Manufacture of office machinery and computers
	31	31.1	Manufacture of electric motors, generators and transformers
		31.2	Manufacture of electricity distribution and control apparatus
		31.3	Manufacture of insulated wire and cable
		31.4	Manufacture of accumulators, primary cells and primary batteries
		31.5	Manufacture of lighting equipment and electric lamps
		31.6	Manufacture of electrical equipment n.e.c.

Table 2.1 (continued); Subsectors, Divisions, Groups and Classes of Sector “D”

PSEK	KVED2	KVED3	Title
DL	32	32.1	Manufacture of electronic valves and tubes and other electronic components
		32.2	Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy
		32.3	Manufacture of television and radio receivers, sound or video recording or reproducing apparatus and associated goods
	33	33.1	Manufacture of medical and surgical equipment and orthopaedic appliances
		33.2	Manufacture of television and radio receivers, sound or video recording or reproducing apparatus and associated goods
		33.3	Manufacture of industrial process control equipment
		33.4	Manufacture of optical instruments and photographic equipment
		33.5	Manufacture of watches and clocks
DM	34	34.1	Manufacture of motor vehicles
		34.2	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers
		34.3	Manufacture of parts and accessories for motor vehicles and their engines
	35	35.1	Building and repairing of ships and boats
		35.2	Manufacture of railway and tramway locomotives and rolling stock
		35.3	Manufacture of aircraft and spacecraft
		35.4	Manufacture of motorcycles and bicycles
		35.5	Manufacture of other transport equipment n.e.c.
DN	36	36.1	Manufacture of furniture
		36.2	Manufacture of jewellery and related articles
		36.3	Manufacture of musical instruments
		36.4	Manufacture of sports goods
		36.5	Manufacture of games and toys
		36.6	Miscellaneous manufacturing n.e.c.
	37	37.1	Recycling of metal waste and scrap
		37.2	Recycling of non-metal waste and scrap

Table 2.2, Sample Composition

	2001	2002	2003	2004	2005	Total
Firms total	280,583	314,711	319,935	333,440	345,909	1,594,578
Manufacturing firms total	44,770	48,151	49,008	49,946	50,719	242,594
Establishments	45,840	49,650	51,205	52,288	53,096	252,079
Positive Labor and Output	35,989	38,040	39,076	38,780	38,634	190,519
Located in QMSA	33,767	35,771	36,790	36,604	36,568	179,500
Urban	25,817	27,519	28,312	28,474	28,739	138,861

Table 2.3, Composition of Manufacturing Sub-Sectors

	2001	2002	2003	2004	2005	Total
DA: Food, beverages and tobacco	6,410	6,365	6,246	5,837	5,530	30,388
DB: Textiles and products	2,705	2,941	3,084	2,937	2,817	14,484
DC: Leather and products	529	553	551	479	441	2,553
DD: Wood and wood products	2,426	2,637	2,697	2,638	2,522	12,920
DE: Pulp, paper, publishing	4,221	4,564	4,746	4,910	5,130	23,571
DF: Coke, refined petroleum products	135	222	224	225	203	1,009
DG: Chemicals, man-made fibres	1,082	1,156	1,180	1,205	1,247	5,870
DH: Rubber and plastic	1,090	1,211	1,337	1,422	1,558	6,618
DI: Other non-metallic mineral products	2,098	2,249	2,349	2,391	2,455	11,542
DJ: Basic metals and fabricated metal	2,382	2,613	2,829	2,996	3,093	13,913
DK: Machinery and equipment n.e.c.	3,918	4,074	4,210	4,158	4,072	20,432
DL: Electrical and optical equipment	3,674	3,920	3,899	3,886	3,929	19,308
DM: Transport equipment	825	900	919	908	948	4,500
DN: Manufacturing n.e.c.	2,272	2,366	2,519	2,612	2,623	12,392
Total	33,767	35,771	36,790	36,604	36,568	179,500

Table 2.4 Composition of Machinery and High-Tech Industry Groups' Samples

	2001	2002	2003	2004	2005	Total
Machinery total	3,042	3,225	3,390	3,375	3,312	16,344
291 (Specialized)	*568*	*584*	*592*	*559*	*564*	*2,867*
292 (General, n.e.c.)	*950*	*1,052*	*1,119*	*1,166*	*1,168*	*5,455*
294 (Machine tools)	*374*	*376*	*408*	*390*	*374*	*1,922*
295 (Other special purpose)	*1,150*	*1,213*	*1,271*	*1,260*	*1,206*	*6,100*
High-Tech total	1,010	1,078	1,023	1,033	1,027	5,171
296 (Weapons)	*18*	*14*	*9*	*17*	*18*	*76*
300 (Computers)	*435*	*451*	*395*	*402*	*380*	*2,063*
321 (Electronics)	*158*	*158*	*170*	*148*	*148*	*782*
331 (Medical equipment)	*338*	*376*	*375*	*384*	*406*	*1,879*
353 (Aircraft and space)	*61*	*79*	*74*	*82*	*75*	*371*

Table 2.5, Descriptive Statistics of the Pooled Data

Year	Employment		Total Assets		Total Sales	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
2001	85	597.0	2,363.5	27,035.7	3,986.5	57,407.5
2002	76	535.4	2,225.5	24,732.1	4,395.0	63,378.3
2003	71	469.3	2,236.5	23,506.4	5,858.7	89,076.0
2004	71	480.0	2,426.1	24,931.9	8,631.7	150,398.1
2005	70	438.6	2,682.6	25,412.4	9,956.2	164,988.1

Table 2.6, Share of Large Firms (Employment >50), by Subsectors

	2001	2002	2003	2004	2005	Subsector Average
DA: Food, beverages and tobacco	29%	28%	27%	29%	31%	29%
DB: Textiles and products	22%	18%	15%	15%	16%	17%
DC: Leather and products	24%	21%	18%	21%	23%	21%
DD: Wood and wood products	7%	7%	7%	7%	7%	7%
DE: Pulp, paper, publishing	7%	7%	7%	7%	8%	7%
DF: Coke, refined petroleum product	35%	22%	25%	23%	24%	25%
DG: Chemicals, man-made fibres	22%	23%	24%	23%	24%	23%
DH: Rubber and plastic	13%	13%	15%	16%	16%	15%
DI: Other non-metallic mineral product	31%	26%	25%	24%	24%	26%
DJ: Basic metals and fabricated metal	21%	19%	18%	19%	19%	19%
DK: Machinery and equipment n.e.c.	26%	24%	21%	20%	20%	22%
DL: Electrical and optical equipment	15%	13%	13%	13%	14%	14%
DM: Transport equipment	33%	31%	30%	32%	31%	31%
DN: Manufacturing n.e.c.	16%	15%	14%	14%	13%	14%
Average	20%	19%	18%	18%	18%	18%

Table 2.7, Shares of Large Firms in Total Employment, Total Assets and Total Output

PSEK	Share of Employment	Share of Capital	Share of Output
DA: Food, beverages and tobacco	87%	90%	91%
DB: Textiles and products	77%	87%	74%
DC: Leather and products	75%	85%	80%
DD: Wood and wood products	50%	68%	72%
DE: Pulp, paper, publishing	55%	83%	70%
DF: Coke, refined petroleum products	97%	96%	92%
DG: Chemicals, man-made fibres	92%	97%	90%
DH: Rubber and plastic	73%	84%	74%
DI: Other non-metallic mineral product	86%	88%	89%
DJ: Basic metals and fabricated metal	93%	97%	97%
DK: Machinery and equipment n.e.c.	90%	94%	86%
DL: Electrical and optical equipment	84%	92%	72%
DM: Transport equipment	96%	98%	96%
DN: Manufacturing n.e.c.	69%	81%	63%
Average	86%	93%	89%

Table 2.8, Means of Major Variables for Large Manufacturing Firms

Subsector	Employment	Capital	Output	Urban
DA: Food, beverages and tobacco	238.9	7,788.8	26,646.9	52%
DB: Textiles and products	256.6	3,912.3	4,228.0	81%
DC: Leather and products	202.9	3,004.5	7,392.5	91%
DD: Wood and wood products	155.7	4,100.5	8,848.0	56%
DE: Pulp, paper, publishing	170.0	7,043.7	16,135.9	91%
DF: Coke, refined petroleum products	1,062.8	93,521.4	443,547.6	87%
DG: Chemicals, man-made fibres	478.3	30,044.9	46,877.7	88%
DH: Rubber and plastic	220.0	11,001.9	20,165.1	82%
DI: Other non-metallic mineral product	265.9	7,857.8	13,450.3	67%
DJ: Basic metals and fabricated metal	737.3	31,486.9	125,235.2	85%
DK: Machinery and equipment n.e.c.	408.2	11,294.1	14,617.1	79%
DL: Electrical and optical equipment	355.6	10,666.8	14,953.3	91%
DM: Transport equipment	793.1	22,515.0	53,507.4	91%
DN: Manufacturing n.e.c.	177.7	3,881.9	13,902.7	82%
Average	345.4	12,006.0	32,107.4	74%

Table 2.9, Means of Major Variables for Small Manufacturing Firms

Subsector	Employment	Capital	Output	Urban
DA: Food, beverages and tobacco	14.9	350.3	1,087.9	56%
DB: Textiles and products	16.2	114.5	310.0	83%
DC: Leather and products	18.2	148.0	515.2	85%
DD: Wood and wood products	11.8	143.5	257.7	64%
DE: Pulp, paper, publishing	10.5	116.5	528.3	85%
DF: Coke, refined petroleum products	12.3	1,507.8	12,717.3	52%
DG: Chemicals, man-made fibres	12.7	294.1	1,640.8	85%
DH: Rubber and plastic	13.7	348.7	1,219.5	86%
DI: Other non-metallic mineral product	15.2	375.8	581.0	66%
DJ: Basic metals and fabricated metal	13.2	255.7	1,018.6	87%
DK: Machinery and equipment n.e.c.	13.7	192.1	657.3	83%
DL: Electrical and optical equipment	10.4	156.1	896.6	89%
DM: Transport equipment	15.5	199.2	907.0	92%
DN: Manufacturing n.e.c.	13.6	157.7	1,402.7	86%
Average	13.2	218.3	862.3	78%

Table 2.10, Means for Major Variables in Machinery and High-Tech Industry Groups

		2001	2002	2003	2004	2005
		Small Firms				
Machinery	Employment	13.2	13.4	13.9	13	12.4
	Capital	137.9	125.7	147.8	183.4	166.3
	Output	451.7	510.2	634	848.9	895
	Urban	90%	89%	88%	90%	90%
	Majority Private	98%	98%	98%	97%	98%
	Foreign Owned	1%	1%	1%	1%	1%
High-Tech	Employment	10.4	10.2	10.9	10.5	10.4
	Capital	114.6	223.2	107.8	117.9	135.4
	Output	486.8	517.5	620.1	1,377.80	1,145.80
	Urban	95%	94%	94%	95%	94%
	Majority Private	97%	97%	97%	98%	97%
	Foreign Owned	1%	2%	2%	2%	1%
		Large Firms				
Machinery	Employment	431.5	424.8	393.3	387.5	399.9
	Capital	9,837.6	10,385.5	10,271.0	10,279.4	11,923.6
	Output	9,973.7	11,114.4	13,114.3	18,484.5	22,202.4
	Urban	85%	86%	87%	88%	88%
	Majority Private	82%	84%	85%	88%	89%
	Foreign Owned	1%	2%	3%	3%	4%
High-Tech	Employment	740.3	614.2	565.5	556.5	493.2
	Capital	20,632.9	16,586.9	16,002.8	15,852.2	15,548.2
	Output	20,151.7	23,119.4	26,403.4	28,952.6	36,967.0
	Urban	89%	91%	89%	91%	90%
	Majority Private	66%	62%	62%	60%	61%
	Foreign Owned	1%	2%	1%	4%	4%

Table 2.11, Average Employment by Ownership Type

	Majority State			Majority Private		
Year	Machinery	High-tech	Year Average	Machinery	High-tech	Year Average
2001	374	884	544	95	61	86
2002	434	573	481	85	52	77
2003	378	568	441	75	47	69
2004	296	617	408	75	43	68
2005	357	395	371	77	45	70
Group Average	369	611	453	81	50	74

Table 2.12, Average Employment in Foreign vs. Domestic Private Firms

	Private, Majority Domestic			Private, Majority Foreign		
Year	Machinery	High-tech	Year Average	Machinery	High-tech	Year Average
2001	93	61	85	203	49	161
2002	79	52	72	510	60	374
2003	69	47	64	511	61	395
2004	68	43	63	488	42	362
2005	70	45	64	436	80	358
Group Average	75	49	69	438	58	337

Table 2.13, Average Capital-Labor Ratios for Urban and Suburban Raions, by Subsectors

Subsector	2001	2002	2003	2004	2005
	Suburban raions				
DA: Food, beverages and tobacco	29.6	31.4	35.9	42.2	35.1
DB: Textiles and products	19.3	16.1	18.5	16.9	15.8
DC: Leather and products	5.5	10	35.5	80.2	66.9
DD: Wood and wood products	11.6	12.5	15.2	22.5	21.5
DE: Pulp, paper, publishing	8.2	8.4	9.9	9.3	14.7
DF: Coke, refined petroleum products	369.5	108.6	47.2	52.5	51.9
DG: Chemicals, man-made fibres	17.9	18	24.2	26.2	33.6
DH: Rubber and plastic	19.6	17.8	44.5	41.7	32.2
DI: Other non-metallic mineral products	24.8	26.1	27.9	23	23.6
DJ: Basic metals and fabricated metal	19	22.3	24.1	49	40.9
DK: Machinery and equipment n.e.c.	20.2	20.6	21.7	25.7	31.5
DL: Electrical and optical equipment	14.9	16	17.8	19	21.3
DM: Transport equipment	131.5	165.7	147.2	30.9	29.6
DN: Manufacturing n.e.c.	8.5	14.4	11.2	22.6	28.2
Group Average	23.4	24.4	26.7	30.2	28.4
	Urban Raions				
DA: Food, beverages and tobacco	15.8	25.8	26.1	29	40.1
DB: Textiles and products	9.1	7.9	8.4	9.7	16
DC: Leather and products	6.6	11.1	11.6	10.4	12.5
DD: Wood and wood products	10.6	11.2	12.3	14.2	23.3
DE: Pulp, paper, publishing	7.9	8.9	11.1	14.4	14.9
DF: Coke, refined petroleum products	59.3	58.8	177.7	121.4	117.4
DG: Chemicals, man-made fibres	21.1	22.1	27.8	41.3	43.6
DH: Rubber and plastic	30.6	27.2	26.1	26	28.5
DI: Other non-metallic mineral products	38.3	33	36	30.5	39.5
DJ: Basic metals and fabricated metal	12.2	13.9	20.1	21.1	26.3
DK: Machinery and equipment n.e.c.	13.3	12.4	15.8	15	16.5
DL: Electrical and optical equipment	11.7	22.9	12.2	13.7	14.2
DM: Transport equipment	14.1	14	15	24.5	19
DN: Manufacturing n.e.c.	9.7	9.9	11.5	14.8	22.5
Group Average	14.2	16.7	17.7	19.3	23.5

Table 2.14, Average Establishment Employment in Urban and Suburban Raions, by Subsectors

Subsector	2001	2002	2003	2004	2005	Sector Average
	Suburban raions					
DA: Food, beverages and tobacco	73	68	64	68	73	69
DB: Textiles and products	54	47	39	39	42	44
DC: Leather and products	36	36	31	30	30	33
DD: Wood and wood products	30	25	24	24	25	26
DE: Pulp, paper, publishing	27	26	24	23	23	25
DF: Coke, refined petroleum products	90	40	59	48	78	57
DG: Chemicals, man-made fibres	52	59	60	44	52	53
DH: Rubber and plastic	32	34	34	36	38	35
DI: Other non-metallic mineral products	101	83	73	74	72	80
DJ: Basic metals and fabricated metal	59	50	49	52	54	53
DK: Machinery and equipment n.e.c.	67	57	46	46	47	53
DL: Electrical and optical equipment	29	27	31	30	30	30
DM: Transport equipment	179	171	128	171	155	160
DN: Manufacturing n.e.c.	46	38	34	35	35	37
Year average	61	54	49	51	53	53
	Urban raions					
DA: Food, beverages and tobacco	84	84	85	92	97	88
DB: Textiles and products	78	64	55	55	52	61
DC: Leather and products	76	60	54	58	60	62
DD: Wood and wood products	19	20	20	20	20	20
DE: Pulp, paper, publishing	22	22	21	21	21	21
DF: Coke, refined petroleum products	495	384	396	442	392	418
DG: Chemicals, man-made fibres	158	142	128	122	114	132
DH: Rubber and plastic	49	46	44	44	44	45
DI: Other non-metallic mineral products	94	81	77	76	74	80
DJ: Basic metals and fabricated metal	209	181	158	153	146	167
DK: Machinery and equipment n.e.c.	135	121	107	101	100	112
DL: Electrical and optical equipment	74	63	59	56	54	61
DM: Transport equipment	312	263	253	278	246	269
DN: Manufacturing n.e.c.	43	38	36	35	35	37
Year average	93	83	77	77	74	80

Table 2.15, Average Firm Total Assets for Urban and Suburban Raions

Subsector	2001	2002	2003	2004	2005	Sector Average
	Suburban raions					
DA: Food, beverages and tobacco	1,709	1,713	1,734	2,087	2,696	1,961
DB: Textiles and products	487	395	330	452	479	425
DC: Leather and products	284	239	292	639	375	357
DD: Wood and wood products	394	406	438	649	817	542
DE: Pulp, paper, publishing	670	774	769	753	826	761
DF: Coke, refined petroleum products	10,304	4,234	5,925	6,766	12,427	7,323
DG: Chemicals, man-made fibres	2,324	2,013	2,176	2,281	4,431	2,698
DH: Rubber and plastic	583	630	1,550	1,088	1,229	1,045
DI: Other non-metallic mineral products	2,193	2,021	1,817	1,993	2,117	2,023
DJ: Basic metals and fabricated metal	1,383	1,203	1,268	1,993	2,836	1,799
DK: Machinery and equipment n.e.c.	1,450	1,338	1,179	1,204	1,317	1,300
DL: Electrical and optical equipment	623	630	973	1,225	1,147	912
DM: Transport equipment	4,025	3,946	3,093	3,897	5,353	4,064
DN: Manufacturing n.e.c.	456	460	382	603	693	522
Year average	1,327	1,285	1,301	1,537	1,931	1,471
	Urban raions					
DA: Food, beverages and tobacco	2,030	2,322	2,730	3,517	4,267	2,939
DB: Textiles and products	978	885	772	773	830	845
DC: Leather and products	862	759	695	821	1,026	824
DD: Wood and wood products	236	283	331	382	501	347
DE: Pulp, paper, publishing	446	481	548	670	745	585
DF: Coke, refined petroleum products	30,035	24,420	37,075	46,402	43,358	36,266
DG: Chemicals, man-made fibres	8,389	8,012	7,615	8,069	7,767	7,962
DH: Rubber and plastic	2,187	2,121	2,034	1,876	2,049	2,045
DI: Other non-metallic mineral products	2,466	2,341	2,346	2,385	2,712	2,452
DJ: Basic metals and fabricated metal	8,122	7,154	6,490	6,404	6,715	6,929
DK: Machinery and equipment n.e.c.	3,397	3,030	2,786	2,750	2,928	2,969
DL: Electrical and optical equipment	1,942	1,665	1,636	1,578	1,555	1,671
DM: Transport equipment	7,864	7,074	6,901	7,442	8,274	7,511
DN: Manufacturing n.e.c.	725	649	742	760	745	726
Year average	2,683	2,508	2,517	2,680	2,887	2,656

Table 2.16, Means of Major Variables for Dnipropetrovsk Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	937	128	3,942.3	8,069.5
2002	1,073	108	3,432.3	7,692.8
2003	1,184	100	3,159.4	9,447.1
2004	1,360	89	2,893.2	10,963.9
2005	1,365	91	3,139.5	14,775.2
Subsector				
DA: Food, beverages and tobacco	802	89	2,717.6	13,604.2
DB: Textiles and products	306	55	757.6	756.8
DC: Leather and products	61	40	337.5	1,865.8
DD: Wood and wood products	242	13	76.9	276.4
DE: Pulp, paper, publishing	702	28	990.2	3,562.5
DF: Coke, refined petroleum products	26	316	15,912.1	39,536.0
DG: Chemicals, man-made fibres	324	56	1,516.0	7,964.5
DH: Rubber and plastic	394	103	4,569.9	13,844.3
DI: Other non-metallic mineral product	292	49	1,400.7	2,530.5
DJ: Basic metals and fabricated metal	699	216	7,854.7	31,122.7
DK: Machinery and equipment n.e.c.	726	91	2,167.0	3,619.2
DL: Electrical and optical equipment	614	61	3,520.2	5,157.4
DM: Transport equipment	158	789	21,357.5	43,189.0
DN: Manufacturing n.e.c.	573	45	841.2	7,122.8
Industry Groups				
Machinery	648	76	1,777.0	3,124.9
High-tech	130	441	12,362.7	8,835.7

Table 2.17 , Means of Major Variables for Donetsk Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	1234	66	2,256.3	5,324.9
2002	1,265	68	2,384.5	8,626.9
2003	1,325	67	2,381.4	13,128.3
2004	1,292	70	2,576.4	25,592.2
2005	1,342	72	2,868.2	18,063.1
Subsector				
DA: Food, beverages and tobacco	746	105	5,840.0	27,381.4
DB: Textiles and products	402	48	1,266.9	1,522.9
DC: Leather and products	68	29	561.1	1,192.2
DD: Wood and wood products	310	12	102.2	359.3
DE: Pulp, paper, publishing	770	19	382.8	1,228.9
DF: Coke, refined petroleum products	31	451	20,009.2	164,273.5
DG: Chemicals, man-made fibres	242	94	3,114.0	14,926.4
DH: Rubber and plastic	297	31	717.2	4,065.8
DI: Other non-metallic mineral product	313	36	2,077.7	2,936.6
DJ: Basic metals and fabricated metal	689	150	6,760.0	61,039.5
DK: Machinery and equipment n.e.c.	1087	107	2,728.3	8,223.7
DL: Electrical and optical equipment	892	31	581.3	4,312.4
DM: Transport equipment	60	52	3,074.0	6,613.6
DN: Manufacturing n.e.c.	551	28	616.8	7,145.3
Industry Groups				
Machinery	1032	71	1,617.5	5,423.1
High-tech	213	49	928.4	3,673.4

Table2.18, Means of Major Variables for Mariupol Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	297	350	14,128.3	33,699.7
2002	298	320	12,484.9	37,753.5
2003	302	293	12,269.2	56,985.8
2004	279	334	14,234.2	99,197.1
2005	305	294	13,939.3	95,316.3
Subsector				
DA: Food, beverages and tobacco	159	102	2,216.1	3,931.7
DB: Textiles and products	114	40	530.1	611.0
DC: Leather and products	13	9	249.8	370.2
DD: Wood and wood products	67	14	192.8	211.6
DE: Pulp, paper, publishing	81	17	402.0	956.0
DF: Coke, refined petroleum products	17	582	49,816.8	222,193.9
DG: Chemicals, man-made fibres	21	128	5,936.7	3,255.2
DH: Rubber and plastic	80	28	1,076.3	1,368.6
DI: Other non-metallic mineral product	81	154	5,980.5	3,699.2
DJ: Basic metals and fabricated metal	192	1609	79,698.8	411,549.5
DK: Machinery and equipment n.e.c.	274	89	1,810.9	3,897.4
DL: Electrical and optical equipment	172	43	963.2	376.5
DM: Transport equipment	118	635	14,182.5	84,090.7
DN: Manufacturing n.e.c.	92	52	2,082.0	1,061.7
Industry Groups				
Machinery	256	79	1,697.9	3,103.1
High-tech	92	17	112.0	348.6

Table 2.19, Means of Major Variables for Zaporizhia Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	838	169	5,725.8	11,893.4
2002	884	155	5,518.8	12,682.8
2003	961	140	5,166.8	16,298.8
2004	971	141	5,623.1	24,973.6
2005	1,006	137	6,272.2	25,176.2
Subsector				
DA: Food, beverages and tobacco	417	106	5,264.7	14,555.2
DB: Textiles and products	189	46	392.9	1,575.1
DC: Leather and products	73	34	202.6	723.1
DD: Wood and wood products	170	23	395.3	631.8
DE: Pulp, paper, publishing	541	15	193.8	815.7
DF: Coke, refined petroleum products	13	1159	62,649.2	395,688.4
DG: Chemicals, man-made fibres	181	93	5,029.1	6,883.8
DH: Rubber and plastic	147	37	980.4	1,565.3
DI: Other non-metallic mineral products	312	120	5,689.9	8,718.3
DJ: Basic metals and fabricated metal	558	398	22,292.7	82,282.2
DK: Machinery and equipment n.e.c.	657	66	1,007.5	2,194.9
DL: Electrical and optical equipment	887	129	3,710.2	5,905.5
DM: Transport equipment	186	844	20,579.6	88,636.6
DN: Manufacturing n.e.c.	329	23	277.7	3,128.5
Industry Groups				
Machinery	596	65	964.9	2,141.9
High-tech	182	485	11,230.8	24,623.2

Table 2.20, Means of Major Variables for Kharkiv Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	2192	75	1,587.9	2,867.8
2002	2,417	66	1,428.1	2,947.7
2003	2,508	62	1,455.9	3,404.0
2004	2,529	59	1,453.5	4,524.6
2005	2,504	59	1,503.7	4,344.1
Subsector				
DA: Food, beverages and tobacco	1043	76	2,906.4	14,613.8
DB: Textiles and products	1458	29	281.7	572.1
DC: Leather and products	191	31	336.1	624.6
DD: Wood and wood products	480	13	95.0	345.8
DE: Pulp, paper, publishing	1599	22	499.2	1,712.9
DF: Coke, refined petroleum products	37	114	4,366.6	15,020.7
DG: Chemicals, man-made fibres	501	66	1,285.3	5,021.3
DH: Rubber and plastic	608	20	575.1	1,747.3
DI: Other non-metallic mineral products	437	60	1,294.3	3,357.9
DJ: Basic metals and fabricated metal	937	35	557.5	1,901.2
DK: Machinery and equipment n.e.c.	1777	141	3,725.7	4,538.2
DL: Electrical and optical equipment	1722	83	1,760.6	3,109.6
DM: Transport equipment	306	239	4,851.0	6,522.2
DN: Manufacturing n.e.c.	1054	29	276.3	2,249.0
Industry Groups				
Machinery	1507	106	2,276.2	3,760.0
High-tech	407	177	3,728.8	4,628.6

Table 2.21, Means of Major Variables for Kyiv Establishments

	Number of Establishments	Average Establishment Employment	Average Capital	Average Output
Year				
2001	4338	45	1,278.1	2,926.7
2002	4,844	40	1,234.6	3,454.1
2003	4,896	39	1,301.0	4,113.0
2004	5,227	39	1,508.2	5,287.2
2005	5,580	37	1,547.5	6,237.1
Subsector				
DA: Food, beverages and tobacco	1636	85	4,232.1	19,647.1
DB: Textiles and products	1588	44	863.4	1,643.2
DC: Leather and products	303	71	1,690.7	4,877.9
DD: Wood and wood products	1343	21	401.9	1,157.5
DE: Pulp, paper, publishing	6784	23	735.6	2,273.4
DF: Coke, refined petroleum products	70	18	1,447.9	34,568.0
DG: Chemicals, man-made fibres	1032	49	2,938.2	8,608.4
DH: Rubber and plastic	1054	36	2,810.2	4,370.3
DI: Other non-metallic mineral products	885	65	2,617.8	7,169.9
DJ: Basic metals and fabricated metal	2093	27	526.7	2,599.5
DK: Machinery and equipment n.e.c.	2384	46	1,037.1	2,634.6
DL: Electrical and optical equipment	3275	37	1,393.5	3,591.7
DM: Transport equipment	586	143	3,897.1	10,736.0
DN: Manufacturing n.e.c.	1852	33	669.1	3,653.2
Industry Groups				
Machinery	2162	40	782.6	2,031.7
High-tech	1379	59	2,075.3	7,308.5

Figures

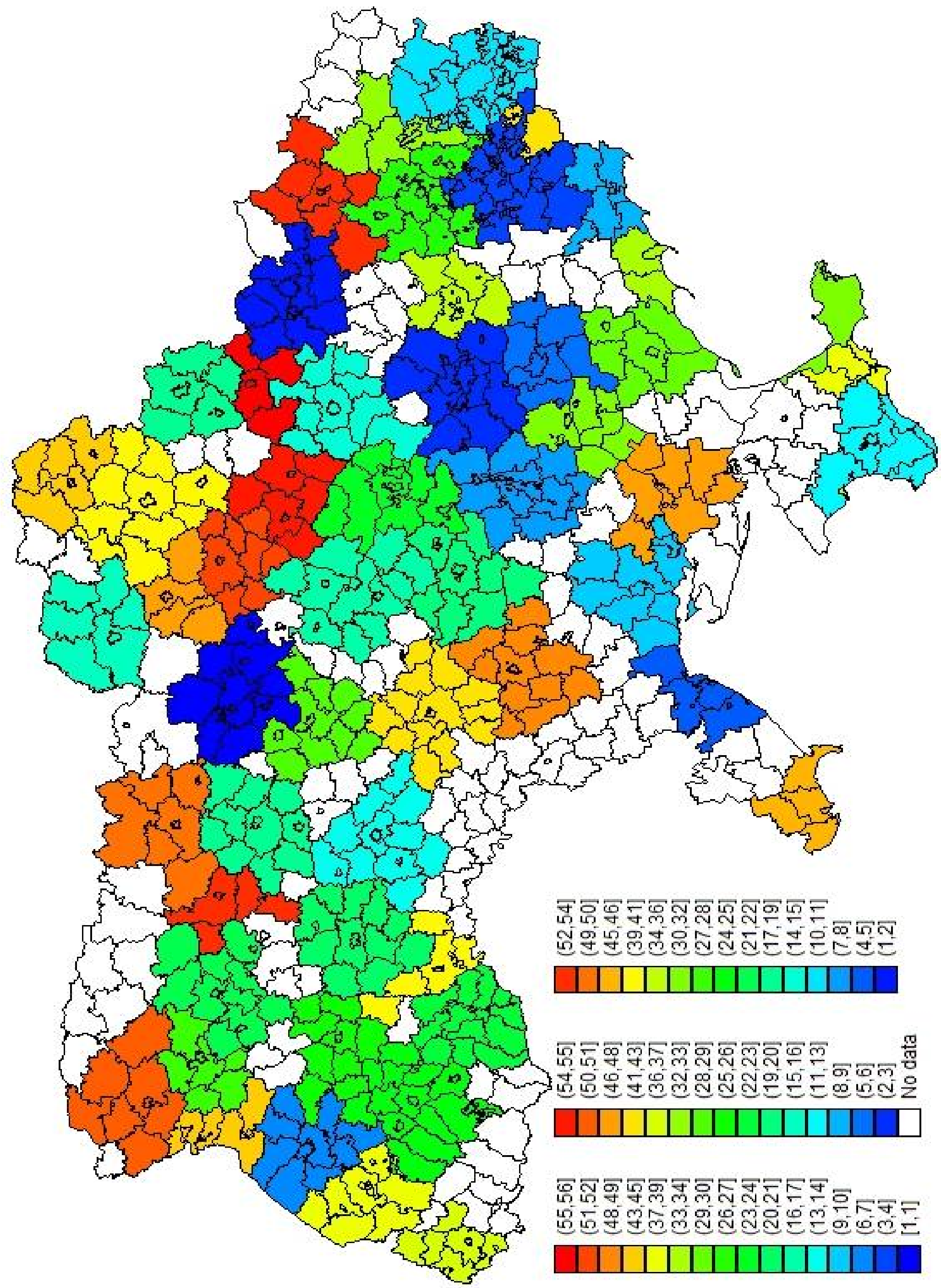


Figure 2.1, Quasi Metropolitan Statistical Areas Based on Sixty Kilometers (about Thirty Seven Miles) Commuting Transportation Distance.

Chapter 3. Spatial Distribution of Manufacturing Industries in Ukraine

3.1. Introduction

In this chapter I explore general patterns in spatial distribution of manufacturing industries in Ukraine. I use plant-level data for period of 2001 – 2005. By the beginning of 2001, the first stage of transition in Ukraine was over. Firms that were once owned by the sate had enough time to adjust their production processes and output lines to the market demands after the privatization and restructuring. The process of adjustment took a significant amount of time. Brown and Earle (2007) found that that the productivity effect of privatization was gained only several years later. In particular, they report that the productivity gap between domestically owned private firms and state firms reached as much as twenty five percent by seven years after the privatization. With caveats regarding the sample size, it appears that for a majority of foreign owned privatized firms, this gap reached forty to sixty percent over the same period. When new firms considered entering the market, they had to take into account the location of existing firms and the market potential of the area. These considerations affected the spatial distribution of manufacturing and determined further development.

Both the population and the total labor force in Ukraine have steadily decreased since 1990. According to the statistical yearbooks (Derzhkomstat (2006)), the total employment has fallen by fifty five percent, from over twenty three million employees in 1990 to eleven million in 2005. More than a half of this fall was registered in the nineties. However, this decrease differed across regions and industries. While the average decrease of total employment in Ukraine since 2000 was seventeen percent, as it is evident from Table 3.1, Central and Western regions experienced a greater fall in employment compared to the Eastern region. Kyiv City was the only region at the first level of the administrative division which experienced a positive growth in total employment at fourteen percent, but the level of employment of 1990 still has not been reached. For a more detailed exposition of the changes in the total employment over the study period, refer to Figure 3.1.

The employment change also varied across sectors of the economy. I present the dynamics of changes in the total employment by sectors in Table 3.2. Agriculture, forestry and fishing combined have lost one half of the employment over the study period. Construction was another sector with a considerable employment drop. At the same time, employment in the financial sector almost doubled. Retail sales and education also experienced more modest increases in employment by four and one percent, respectively.

As I show in Table 3.3, the decline of manufacturing employment (excluding raw materials extraction and energy distribution facilities) between 2000 and 2005 has reached nineteen percent over the study period and was close to the economy average. On the other hand, the change in employment was also different between sectors. In three manufacturing subsectors, DE ("Pulp, paper, publishing"), DF ("Coke, refined petroleum products"), and DH ("Rubber and plastic"), employment increased by nine to fourteen percent. In all other subsectors, employment fell. Textile and leather production were the two subsectors where employment decreased the most, by forty three and forty nine percent respectively. Machinery (subsectors DK ("Machinery and equipment n.e.c."), DL ("Electrical and optical equipment"), and ("DM: Transport equipment") was another group of industries that experienced a significant employment decline of twenty seven percent combined.

The structure of the economic activity seems to have changed over the last fifteen years. Even though the official employment statistics does not include people employed by "statistically small enterprises" and private entrepreneurs, which were absent in the Soviet statistics in 1990, as well as military personnel, which used to be included before 2004, the contraction of the employment by the year 2005 still appears to be striking. First of all, a drastic decline in agricultural employment may be explained by the dismissal of collective farms where employment was obligatory during the Soviet times, and the following relocation of some rural residents to cities. Since the Western and the Central part of Ukraine are predominantly rural, this fact may explain a more than proportional employment decline there. On the other hand, many urban residents with higher human capital could have leaved manufacturing where the output and wages fell during the nineties, and chose occupations in retail trade which demonstrated significant growth. A two-fold increase in the financial sector employment may be connected with two possible factors. First, the sector has been expanding territorially into smaller towns and rural areas, and second, jobs in this sector paid the highest wages in the economy.

Both factors attract younger cohorts of employees that came into the labor market after 2000. Finally, natural population decline (especially pronounced in rural areas) and migration outflows of qualified personnel (with border oblasts leading in this phenomenon) may also contribute to the economy-wide decline in total employment and ultimate relocation of labor force both between regions and between sectors. As a result of restructuring, some sectors became relatively more concentrated in certain regions over time.

In the reminder of the chapter I describe patterns of industrial concentration and agglomeration in Ukrainian manufacturing. In the next section, I explore changes in manufacturing subsectors using choropleth maps, regional diagrams and location quotients. In the third section, I discuss changes in machinery and higher technology industry groups, as I concentrate on these two sectors in the next chapter. In the fourth section, I present an analysis of Ellison-Glaeser type agglomeration coefficients. The fifth section concludes.

3.2. General Outlook of Spatial Distribution in Manufacturing

Particularities of industrial landscape in Ukraine were initially determined by two factors. First, abundance of raw materials (ores, energetic coals, and major rivers suitable for sailing and construction of hydro electric plants) led to a development of metallurgy, coal mining, and various types of machinery in the Eastern part of the country since the middle of the nineteenth century. On the other hand, central and Southern regions endowed with rich soils and long periods of insolation became natural domains for agriculture-related industries, such as food processing and textile. Second, Ukraine has historically seated at the crossing of several important transit corridors, connecting Russia with the Western Europe and Baltic states with countries of the Southern Europe and the Middle East. Cities located along those trade routes (such as Lviv, Kyiv, Kharkiv, and Odessa) became important centers of retail trade and knowledge exchange.

The policy of massive industrialization during the Soviet times strengthened disparities in the industrial structure between economic regions of Ukraine. The Eastern part has been increasingly industrialized, with the developed transportation network, especially railroads. For example, the sorting railway station Yasynuvata in Donetsk region is recognized as the largest in the former USSR and one of the largest in Europe.

The region is heavily urbanized, with over eighty percent of urban population. The region is located on the border with Russia, which determined export flows and general cooperation with Russian firms. The Eastern region is abundantly endowed with natural resources, with coal, iron and magnesium ores, salt, and mercury being the most important as the base for the manufacturing.

The Central region is abundant with rich soils and forests, as well as with peats, brown coals, natural construction materials. A large part of the region cannot be used in economic activity due to radioactivity contamination after Chernobyl accident. The city of Kyiv is the largest in the country, and Kyiv agglomeration is the most significant economic center of the region, with all sectors represented in the structure of manufacturing. At the same time, agriculture, food processing and services form the mainframe of the regional economy.

The Western region has lower density of population, which is predominantly rural. Lviv is the only large city. The region is split by Carpathian Mountains. The natural resources, such as multiple sources of minral water, are most suitable for agriculture and tourism, but mineral resources (brown coals, oil, natural gas, polymetalic ores, sulphur and stones used in construction) are also available. Due to its location, the region plays an important role in the transit of the natural gas from Russia to the Western Europe, hosting several large underground storages for natural gas.

The Southern region is developed mostly along the sea costal line. The region has traditionally hosted ship building enterprises, a large amount of food processing plants, and substantial recreation resources.

Ukraine has a well developed energy-generating network. Approximately half of electricity is produced by four nuclear power plants, with another half produced by thermal power plants and less than five percent share of hydro power plants. The share of nuclear plants in the total energy production has doubled over the last fifteen years, mostly due to the decline of the thermal sources. At the same time, Ukrainian manufacturing must rely on imported natural gas and oil, since domestic sources do not cover own needs.

The transportation system in Ukraine is also developed, with automobile transport and railroads playing the leading roles. The Eastern region, Kyiv agglomeration and Lviv oblasts have the highest density of railroads. On the other hand, only 35% of railroads are electrified (Smal, 2004). Kyiv, Kharkiv, Lviv, Dnipropetrovsk and Zaporizhia are the most important knots in the automobile roads network, tough the average density of the automobile roads is lower compared to most European countries. Water and air transportation are not that developed and play an auxiliary role after a drastic decline during the transition period.

Since absolute majority of those firms were built around cities, industrialization was followed by a rapid urbanization of the area. Not surprisingly, the spatial distribution of economic activity in Ukraine is highly correlated with the distribution of population. Most economic activity is concentrated in cities. Table 3.4 presents the number of plants and employment in urban areas. The share of all plants in urban areas has increased from seventy five percent in 2001 to seventy seven percent in 2005. The share of total urban employment increased from sixty six percent in 2001 to seventy four percent in 2005. In manufacturing, the share of plants in cities increased from seventy three percent to seventy five percent, whereas the share of urban employment increased from eighty percent in 2001 to eighty two percent in 2005. These calculations show that both businesses and employment were increasingly concentrating in cities which contributed to greater geographic concentration of economic activity in general and manufacturing in particular. At the same time, many urban settlements have not developed into full scale cities, but rather remained at the status of satellite settlements by larger cities or factory-towns. According to Balabanov (2003), there currently are at least three hundred such settlements in Ukraine.

To illustrate these patterns I constructed quintile choropleth maps of manufacturing employment and number of plants in 2001 and 2005. To show the trend, in both years the quintiles of the year 2001 were used. The employment map is shown at Figure 3.2, and the spatial distribution of plants is at Figure 3.3. The shading increases with the change in employment in that raion. Smaller polygons show cities. To capture the scope of changes I created an indicator variable which is equal zero if the five year change of the characteristic in the raion is within ten percent of the 2001 value, negative one if the characteristic declined by more than ten percent, and one if the variable increased by more than ten percent. I mapped this variable for manufacturing employment and the number of plants, as shown at Figure 3.4. It appears that significant increase in the local raion employment took place predominantly in big cities, such as Kharkiv, Donetsk or Lviv, or in adjacent suburban raions. In the absolute majority of rural raions there was a decline in manufacturing employment. Despite the total growth in the number of plants in the economy, in more than a half of rural raions this number actually fell. It is remarkable that even though the total employment in Kyiv has increased (as it is evident from Table 3.1) manufacturing employment did not significantly change, unlike the number of plants. It is possible to conclude that considerable relocation of the labor force from rural to urban areas took place during the period of study. At the same time, the number of firms has increased in a much greater number of raions, which contributed to the decline of the average plant size.

The quintile points clearly indicate unequal distribution of economic activity between raions. To illustrate this inequality I have built a series of Lorenz curves for the distribution of employment and plants across raions. The Lorenz curve shows a degree of inequality in a characteristic's actual distribution in comparison with the hypothetical (uniform) distribution of that characteristic. The area between the forty-five degree line and the Lorenz curve is Gini coefficient and represents the numerical measure of such an inequality. The greater is the deviation of the Lorenz curve from the forty five degree line, the greater is the inequality of the distribution of the characteristic. The employment distribution is shown at Figure 3.5, whereas the plant distribution is at Figure 3.6. Since the both employment and the number of plants are highly correlated with the population (with the correlation coefficients of 0.9 and 0.82, respectively), I decided to use unweighted variables. The lower part of each graph is an excerpt to zoom out the difference between years. In general, the Lorenz curves show increasing inequality in the spatial distribution of manufacturing employment and firms. Several big regions, such as Kyiv and other cities with population over 1 million people, are very different from the rest of raions to ensure such a substantial inequality. Since they keep attracting new firms and employment at the expense of other raions, the inequality is also increasing over time.

Another tool in my analysis is the test of spatial autocorrelation. The test statistic is Moran's I (as found in Anselin (1988)). Lafourcade and Mion (2007) propose an intuitive interpretation of Moran's I in the regression context. Let us denote economic activity measure with an X and use a spatial weight matrix w. One can build a *spatial lag* for each raion, wX, by calculating a weighted sum of the economic activity around each region. The slope coefficient in the regression of the spatial lag wX on X is Moran's I. As Lafourcade and Mion (2007) indicate, Moran's I is the "correlation coefficient between X_i and that of its neighbors." If neighboring areas are similar and the number of such similarities is sufficiently large, I reject the null hypothesis that economic activity is randomly distributed over the space.

I used a simple adjacency criterion to build a spatial weight matrix, where each element is zero or one depending on whether two raions have a common border. The diagonal elements are always zeros. As Anselin (1988) suggested, the matrix is row-standardized: each row is divided by the sum of its elements. Kyiv (a capital city) is a distinct outlier where the number of firms exceeds fifteen standard deviations from the average number of firms in a raion. To address this, I experimented with another spatial weight matrix w^* from which Kyiv was excluded. I present the test results in Table 3.5. The normalized value of the test statistic increased from 0.976 in 2001 to 1.451 in 2005 when employment measure was used, whereas for the plant measure, the test statistics increased from 1.380 in 2001 to 1.764 in 2005. One can see that even though I fail to reject the null hypothesis about the randomness of the spatial distribution both in the number of manufacturing plants and employment in all cases but the number of plants in 2004 and 2005, the test statistics tends to increase with time. Also, exclusion of Kyiv did not affect the test statistics much. I believe that the major reason for such behavior of the test statistic is the structure of the spatial weight matrix that has a substantial number of regions with a single neighbor, when a city is completely enclosed within a raion. Once firms tend to concentrate predominantly in cities the spatial lag for most of their neighboring raions will be small. A steady increase in the test statistics may indicate that cities become "overcrowded" or too expensive for running business over time, and firms start seeking for new locations in neighboring rural raions, thereby increasing the spatial correlation. This hypothesis may be supported by the visual presentation of changes in the number of plants and employment in Figure 3.4, which indicated a tendency to grow both in cities and suburban raions, especially near large cities.

I have also constructed diagrams of the distribution of economic activity among subsectors by oblasts. The plant measure is shown at Figure 3.7, whereas employment measure is presented at Figure 3.8. From the diagrams it is difficult to conclude that the structure of manufacturing within each region significantly changed. Even if changes were present, they are subtle, mainly because the time span is rather short. Neverthelss, there is a clearly visible difference between the shares of employment and the shares of plants in some subsectors and similarity in others. For example, employment in manufacturing of basic and fabricated metals (subsector DJ) takes almost one third of the total employment in three oblasts in the Eastern part of the country, in Dnipropetrivsk, Donetsk, and Zaporizhia. However, the share in the number of plants in this subsector in the same three oblasts is approximately one half of the employment measure. A similar observation is valid for the difference between employment and plants shares in manufacturing of machinery and equipment (subsector DK) in Sumy oblast, where employment constitutes about two fifth of the regional manufacturing employment, whereas less than a quarter of plants are involved into this activity. This pattern indicates a presence of a small number of large firms, with the average employment greater than in the entire manufacturing sector. When this happens, a plant-based measure of localization economies is lower than employment-based measure for the particular region. In this respect, the arguments put forward by Holmes and Stevens (2002) and developed in Lafourcade and Mion (2007) about the effect of the plant size on concentration and agglomeration should be taken into account. Holmes and Stevens (2002) found that plants located in areas with larger industry concentration tend to be larger in size, especially in manufacturing. Lafourcade and Mion (2007) used Italian microdata and found that large plants tend to be relatively more spatially concentrated compared to the small plants.

To explore relative intensity of particular subsectors in manufacturing in regions, I calculated location quotients. The location quotient, LQ, is a ratio of the industry share in the region to the region's share in the total employment or plants count in the economy. The measure indicates a relative specialization of a region in the industry. Low values of the quotient (below one) indicate underrepresentation of the industry in the region. When the value of LQ is above one, it means that the industry is overrepresented in the region relative to the entire economy. O'Donoghue and Gleave (2004) mention that areas with high levels of LQ may constitute clusters of economic activity, since they indicate an above average geographical concentration of employment or the number of plants in that industry. However, they also point to a disagreement in the literature about the threshold value of LQ to consider it as a potential indicative of clustering. On the other hand, if the total number of observations within an industry is small compared to the total number of regions, it is possible to get very high levels of LQ. While in some cases such high values may indeed suggest spatial clustering, it is equally possible that activity is just sparsely distributed in a few detached regions without actual formation of regional clusters.

I calculated the location quotients for every raion and every year at the level of two-letter manufacturing subsectors. In Table 3.7 I show the number of raions each subsector has at least one plant. The total number of regions in Ukraine per year is 669, and we can see that no subsector is present in all raions. To explore the general patterns in location quotients across industries and over time, I grouped raions into those where LQ is above one and those where LQ is below one. As we notice from Table 3.8, there are only two industries, DA (food, beverages, and tobacco) and DC (leather and leather products) where the number of raions where this industry dominates exceeds the number of raions where the industry is underrepresented. This pattern does not change over time. While in the case of the subsector DC the difference is minor, in the case of subsector DA, the difference is more pronounced. However, food production is often considered a local industry which supplies a great share of its output to the home market, which may explain relative representativeness of this industry in many raions. Other than that, employment in other manufacturing subsectors is relatively concentrated only in a number of raions which may suggest localization effects. A remarkable example is a subsector DJ (other non-metallic mineral products) where the share of raions where this industry is overrepresented is below twenty percent. This may indicate concentration of this subsector only in a small number of regions. Indeed, as it is evident from Figure 3.8, this industry constitutes a major share of employment in only three oblasts. Considering raions with very large values of the employment-based location quotient in any subsector (LQ>10), we notice that on average, they constitute less than five percent of all raions when a particular subsector is overrepresented, which may also suggest significant geographic concentration of industries in these raions.

In Table 3.9 I present the grouping of the location quotients based on the number of plants in the raion. If we compare these numbers with those in Table 3.8 we notice that the number of raions where a given subsector is relatively underrepresented is greater when I use employment-based location quotient rather than the corresponding plant-based one. This is especially evident for the subsector DF (coke, refined petroleum products and nuclear fuel) which is relatively underrepresented in about two thirds of all raions if measured by employment, but only in ten to thirteen percent of raions, if measured by the number of plants. Another similar case is that of the subsectors DG (chemicals, chemical products and man-made fibers) and DM (transport equipment): it is relatively underrepresented in employment in about two thirds of the raions, but only in about a one third of the raions if I measure LQ by the number of plants. From Table 3.6 one can see that these three subsectors have the greatest average employment per plant. Since these three subsectors have a small number of plants in a small number of raions, such pattern may indicate both high geographic and industrial concentration in these three subsectors.

Analysis of choropleth maps, Lorenz curves, tests for spatial autocorrelation and location quotients suggest that spatial distribution of manufacturing activity in Ukraine is far from being homogenous. While both the total employment and manufacturing employment fell over the period from 2001 to 2005, some regions enjoyed an increase in local employment, which suggests geographic relocation of production factors. The same refers to the number of plants. Even though the total number of plants increased, it fell in some regions and increased more than proportionally in others. Manufacturing subsectors also do not evenly spread across the economy. Location quotient analysis suggests that there may be several spatial clusters in various industries. Spatial correlation shows a slight tendency to increase over time in all manufacturing subsectors, although the test results are not strongly conclusive, and further analysis of spatial clusters may be required. The time period is quite short to ensure much variation in the data, but the exploratory spatial data analysis lets us to make preliminary conclusions about a possibility of agglomeration effects in at least several industries and their intensification over time.

3.3. Trends in the Spatial Distribution of Machinery and Higher Technology Industry Groups

In this section I focus upon two industry groups, machinery and high tech, since I am going to perform a thorough analysis of localization effects in these two industries in the next chapter. Following Henderson (2003) I define *machinery* as the group of three digit industries according to NACE/KVED classification: 291 (Manufacture of machinery for the production and use of mechanical power, except aircraft, vehicle and cycle engines), 292 (Manufacture of other general purpose machinery), 294 (Manufacture of machine tools), and 295 (Manufacture of other special purpose machinery). The *high-tech* industry group consists of the following industries: 296 (Manufacture of weapons and ammunition), 300 (Manufacture of office machinery and computers), 321 (Manufacture of electronic valves and tubes and other electronic components), 331 (Manufacture of medical and surgical equipment and orthopedic appliances), and 353 (Manufacture of aircraft and spacecraft).

Table 3.10 presents general descriptive statistics. While employment in both groups decreases between 2001 and 2005, the total number of plants increases. The remaining plants become smaller in size, which decreases industrial concentration. On the other hand, the number of raions where both industry groups have their plants also diminished by 2005. If we consider maps (see Figure 3.9 and Figure 3.10 for machinery and Figure 3.12 and Figure 3.13 for high tech), we can notice that both industry groups are quite concentrated spatially. Half of all machinery plants are located in only twelve raions (actually, all these raions are large cities in the eastern part of the country), with six raions having more than 100 machinery plants in each. On the other hand, of all raions where machinery is present, about forty percent (140 raions, on average) have only one plant in that activity. The high-tech industry group is also spatially concentrated. One half of all high-tech plants are located in only six large cities, and the city of Kyiv is the only raion that holds more than one hundred high-tech plants. Such inequality of the spatial distribution in both industry groups suggests that the possibilities for localization economies are limited for plants located in single-plant raions, and the most agglomeration effects are visible only in a few raions.

The global Moran's I test (see Table 3.11) supports such a spatial pattern pointing to the absence of non-random spatial correlation for both plants and employment measures. The value of the test statistic is also decreasing with time. I believe the major reason for such behavior of the test is a large number of raions with zero activity in a group, which are difficult to omit given a particular structure of the spatial weight matrix, whereas the test by design treats them as adjacent regions with the same (zero) value of activity. The number of clusters where activity spreads over raion borders (and hence contributes to the spatial correlation measure) is relatively small, which may decrease the value of the test statistic. On the other hand, the pattern of the spatial distribution of economic activity in both industry groups may require additional analysis of local clusters, such as several largest cities where most activity is located.

If we refer to the ownership structure of both industry groups (see Table 3.12), we notice that both machinery and high tech are represented by almost entirely private firms. The total share of private firms in machinery reaches ninety six percent in 2005, whereas in the high tech group, ninety two percent of all plants are private. The share of majority foreign plants is the same in both groups and reaches its maximum at the level of two percent in 2005. However, the picture is different if we refer to employment. The share of private sector employment in machinery varies between seventy eight and eighty four percent, while the share of employment in foreign owned plants increases sharply from 2.4% in 2001 to 9.6% in 2005. In the high tech group, the share of employment at private plants changes from forty two percent in 2001 to fifty seven percent in 2005, whereas employment shares at foreign owned firms increases from 0.6% to 2.3% in 2005. This implies, first, that privatization is still not complete in a few big high tech plants, and second, big machinery plants are more attractive for foreign owners.

Now I turn to the spatial distribution of ownership in both industry groups. We see from Table 3.13 that both groups became more spatially concentrated: while the total number of plants increased, the number of regions where they are located fell. At the same time, foreign ownership spread out in machinery and virtually did not change in high tech: the number of raions with foreign owned machinery firms doubled to forty two in 2005, while in the high tech group, it increased form nine to twelve. On the other hand, most foreign owned firms preferred to locate in the capital city: in 2005, Kyiv hosted thirteen out of sixty eight plants in machinery and eight out of twenty one plants in high tech. Lviv in the West and Odessa in the South also hosted three to five foreign firms in machinery. In fifteen other raions where foreign firms were present in either industry group, there were two of them, and in the rest of raions there was only one. Therefore, there is no case of spatial concentration of foreign ownership in Ukraine except one "hot spot" in the capital of the country where the share of foreign firms is relatively higher compared to other raions.

The number of state-owned firms in each raion did not substantially change over time. Kyiv, Kharkiv, Donetsk and Odessa, the four largest cities in Ukraine, host between eight and fourteen state owned machinery plants each. Kyiv, Kharkiv and Lviv are also cities with the greatest number the state owned plants in the high tech group. However, no single raion or even oblast has a significant concentration of state owned firms. Therefore I may claim that sate ownership in these two sectors is not spatially concentrated, again, except several "hot spots".

While both machinery and high tech industry groups are geographically concentrated in a few raions, mainly in big cities, the ownership patterns are not that clear. The absolute majority of plants in both groups are private domestically owned. A small group of private foreign plants is spread out across the country, with a relatively large number of them in Kyiv, compared to other raions. A larger group of state owned plants is also spread out across many raions, with minor concentration in several big cities. Nonetheless, it is not correct to assume that either foreign owned plants or state owned plants are geographically concentrated to significantly affect localization economies in any raion.

3.4. Indices of geographic concentration

We have already witnessed that manufacturing activity in Ukraine is distributed heterogeneously in space, and the spatial inequality tends to increase over time. In this section I will follow the methodology proposed by Ellison and Glaeser (1997) and explore how random this heterogeneity is. In other words, I will calculate agglomeration indices which measure agglomeration within an industry as an excess of geographic concentration "over and above of that which would be expected given the extent of industrial concentration in the industry" (Devereux et at. 2004)). Unlike previous measures of geographic concentration such as Location Quotient or Gini index, for which treated the "benchmark" case is equal distribution of economic activity across all regions, Ellison-Glaeser agglomeration index compares the geographic concentration of an industry with an expected value based on the industrial concentration and general distribution of economic activity across the economy.

For an economy which consists of *M* regions, the index of agglomeration γ in every industry is calculated as:

$$\gamma = \frac{G - H}{1 - H},$$

where *H* is the Herfindahl index of industrial concentration, $H = \sum_{j=1}^{N} z_j$ (z_j is the share of a plant *j* in the industry employment), and *G* is the Gini-type index of "raw" geographic concentration:

$$G = \frac{\sum_{i=1}^{M} (s_i - x_i)^2}{1 - \sum_{i=1}^{M} x_i^{\,2}}.$$

Here, s_i is the share of the industry employment in the region *i*, and x_i is the region's *i* share of total manufacturing employment (*i* = 1…*M*).

Maurel and Sedilot (1999) developed an agglomeration index very similar to that of Ellison and Glaeser (1997). The difference is in the way of calculating *G*. In Ellison and Glaser (1997), $G^*_{EG} = \sum_{i=1}^{M} (s_i - x_i)^2$ (I omit the correction factor $1 - \sum_{i=1}^{M} x_i^{\,2}$ for the ease of argument). Maurel and Sedilot (1999) proposed to calculate $G^*_{MS} = \sum_{i=1}^{M} s_i^2 - \sum_{i=1}^{M} x_i^2$. Both measures, G^*_{EG} and G^*_{MS} are indicators of the "raw" concentration of an industry, and as a result, both indices are unbiased estimators of a "true" agglomeration of an industry which follows from a theoretical model. However, there is a difference in the inference. Mare (2005) provides the following argument about this difference:

The MS index indicates whether the profile of shares (s_i) for an industry [as shown by a histogram of distribution of the industry employment between regions] *is steeper than the profile of shares (x_i) for total employment. A steeper profile is represented by a positive index; a flatter profile by a negative index. In contrast, the EG index measures how well the profile of industry employment matches that of total employment, with both steeper and flatter profiles being represented by a positive index. The EG index can be negative only as a result of the influence of the Herfindahl index.*

In this paper I use Ellison and Glaeser's (1997) version of the agglomeration index, because I am interested in the distribution of activities across regions. It is not clear how to interpret specific values of this index. Ellison and Glaeser (1997) proposed the following threshold values: $\gamma < 0.02$ imply small concentration, if $0.02 \leq \gamma \leq 0.05$ such an industry is considered "moderately localized", or medium concentration, whereas values of $\gamma > 0.05$ indicate high concentration. Negative values of γ deserve special attention as they indicate cases when spatial distribution of plants across regions is lower than predicted by the industry concentration. Ellison and Glaeser (1997) proposed that such industries may need to be near final consumers, and that is why they are located more sparsely than manufacturing employment on average. However, another reason for negative values may follow from excessive industrial concentration within an industry and the fact that an industry is located in just a few regions.

I calculate the index of agglomeration, γ, for three-digit industries at various levels of regional aggregation. First I calculate the index for all 669 raions. The summary results are present in Table 3.14. Both mean and median values of the index seem to increase over the five year period. At the same time, the share of industries with negative values of the index is almost one quarter, which is surprisingly high compared to other studies. For example, in Ellison and Glaseser (1997), there were only 13 *four-digit* industries (out of 459) with negative concentration. About a half of industries show small concentration, and this total number does not change over time, whereas the number of moderately and substantially concentrated firms slowly increases. I present histograms for 2001 and 2005 in Figure 3.15 to illustrate frequency distribution of gammas with interval widths of 0.01. In both years, distributions appear to be skewed, with the mean being larger than median. The patterns seem similar, but in 2005, the longest bar which corresponds to the number of least concentrated industries ($0 < \gamma < 0.02$) is shorter compared to 2001, whereas the right tail stretches further into "highly concentrated" range. This general result is in line with the observation from the previous section that average geographic concentration of industries in Ukraine tends to increase, and manufacturing seems to drift towards larger cities.

In Table 3.15 I show the distribution of three digit industries by subsectors. The last column for each year is the average plant size. The most industries with high levels of concentration are observed in the subsector DE (KVED 21-22, "pulp, paper and paper products; publishing and printing"), whereas industries in the subsector DK (KVED 29, "machinery and equipment") appear to be least concentrated.

In Table 3.17 I present ten most concentrated and ten least concentrated industries in 2001 and 2005. From the ten most concentrated industries in 2001, six remained on the same list by 2005. Among the ten least concentrated industries in 2001, eight remained on the list by 2005. However, the relative position of industries within each group has changed slightly. For example, "Publishing" (KVED 221) was the most concentrated industry in 2001, but fell to the sixth position in 2005 with twofold drop in the value of the index. On the other hand, "Manufacturing of musical instruments" (KVED 363) was at the ninth place in 2001, but ranked second in 2005. The value of the index depends on the relative size of the industry and the number of regions where the industry is located. In the case of musical instruments, the number of plants dropped one third, thereby increasing the "raw" concentration, and this could have affected the index more compared to an increase in the industrial concentration. On the other side of the spectrum, the relative position and the composition of the group of the least concentrated industries have not significantly changed. In these industries, raw geographic concentration is most often offset by large industrial concentration, which drives the total value of the index into the negative range. It appears that industrial concentration in industries with negative agglomeration coefficient prevails over geographic raw concentrations, which is supported by relatively larger size of plants and a small number of regions where the industry is located.

Holmes and Stevens (2002) found that geographic concentration may be correlated with the plants size, though the direction of causality has yet to be established.

Lafourcade and Mion (2007) argued that EG index neglects this possible correlation and proposed to calculate the "unweighted" version of the index where the measure of economic activity is not the employment in the industry-region cell, but the number of plants instead. Actually, the plant-based ("unweighted") estimator for agglomeration index was proposed in the model of Maurel and Sedilot (1999), hence Lafourcade and Mion (2007) complemented EG model with an "unweighted" version to explore concentration in Italian manufacturing industries.

Since manufacturing employment in Ukraine does not seem to be distributed proportionally to the number of plants in each industry-region cell in several industries, I decided to calculate the unweighted (plant counts based) version of the Ellison-Glaeser agglomeration index as well. The summary of the results is presented in Table 3.16. Comparing with the employment-weighted version of the index in Table 3.14, I would like to emphasize the following issues. First, mean values for the unweighted index are generally smaller, and the entire index is more skewed to the center of the distribution. Second, the mean value is also increasing by 2005, though not as much as in the weighted version. Plotting the values for all industries on the histogram (see Figure 3.16) we notice that in 2005, the distribution of the index values looks flatter and has a thicker right tail compared to 2001. In general, I may conclude that both indices show a slow tendency of concentration in manufacturing industries to increase over time. I also list the ten most and least concentrated industries in Table 3.18. Comparison of employment based rankings and plant count based rankings reveals that in 2001, three industries (KVED 267 "Cutting, shaping and finishing of stone", KVED 272 "Manufacture of tubes", and KVED 351 "Building and repairing of ships and boats") are found among the top ten industries for both versions of the index, whereas in 2005, the same five out of ten industries are considered "most concentrated" despite the version of the index is used. Spearman rank correlation between weighted and unweighted indices is 0.36 in 2001 and 0.46 in 2005, and in both cases the hypothesis of rank independence is rejected. Since the index is sensitive to outliers both in the size of firms and the number of regions hosting an industry, such an increase in the rank correlation may indicate a certain "leveling out" of industries.

Comparing the ranks of the most concentrated industries between 2001 and 2005, as measured by the unweighted index, we also see that same six industries are included in both lists. Spearman rank correlation between estimates of the index in 2001 and 2005 is 0.66 for the unweighted version and 0.65 for the weighted version. In general I may conclude that there are no evident advantages of using the unweighted version compared to a more traditional weighted one.

Finally, since the index is calculated in the same manner as the one in Lafourcade and Mion (2007), we have a rare opportunity to run a cross-country comparison. Even though there already are a number of papers where agglomeration indices are calculated for various countries, usually the values of indices are difficult to compare directly due to differences in the data quality and levels of aggregation, the particular method used, or systems of industry classification. In my case, these differences are minimal, since I also use plant-level data, the same NACE-based industry classification and the same unweighted Ellison-Glaeser index. Even the number of regional units is comparable: 669 raions in Ukraine vs. 784 LLS units in Italy. Lafourcade and Mion (2007) split all firms into small and large ones, having chosen 20 employees as a threshold value. In the case of Ukraine I may assume that most industries are dominated by large firms if I use the same criterion. The average size of a Ukrainian firm is still larger than of a European counterpart, despite the tendency to downsizing since the Soviet times. That is why I compare Ukrainian firms with the "large firms" part of Lafourcade and Mion (2007). I refer to the list of ten most concentrated firms Ukrainian firms as presented in the upper part of Table 3.18 and Table 5 in Lafourcade and Mion (2007). We see that while in 2001 there was no resemblance, in 2005 two industries (KVED 221 "Publishing" and KVED 223 "Reproduction of recorded media") are found among the ten most concentrated both in Ukraine and Italy, and the latter industry ranks first in both countries! If we compare the ranks of the most concentrated industries as measured by the employment-weighted index, we find four same industries in the lists both countries. It may suggest that at least the structure of mass production activities in the Ukrainian economy slowly converges to the market economies of the West.

3.5. Conclusion

Analysis of plant-level manufacturing data in Ukraine over 2001-2005 indicates substantial spatial heterogeneity in virtually all industries. About seventy five percent of all plants and eighty percent of employment in manufacturing are located in cities and towns. Both population and employment in Ukraine fell in the study period. On the other hand, not all industries and not all regions indicate an equal drop in employment. Even though I had to reject a hypothesis of spatial correlation of employment between regions, I also noticed an evident upward trend as rural firms relocate closer to cities, whereas new firms sometimes chose locations in rural raions surrounding cities. The average size of firms continues to drop every year. I tend to believe that the main reason for this is an increasing number of new firm births rather than massive downsizing. Large firms (those with over fifty employees) constituted eighteen percent of all firms in manufacturing but provided work for eighty five percent of the total employment in 2005. These numbers imply that most large Ukrainian manufacturing firms are concentrated disproportionally in several big cities, with the rest of the country being virtually unpopulated in terms of presence of manufacturing. Relocation of employment towards urban centers increases the spatial inequality. This issue becomes especially evident at lower levels of data aggregation. There are approximately 11,300 three-digit manufacturing industry-raion cells in my data. In fifty seven percent of them there is only one firm or a plant. It means that this stand-alone establishment is not able to enjoy localization economies at the most local level. On the other hand, those several locations where most firms are concentrated provide lucrative opportunities for development due to increasing number of spillovers between firms. My data analysis confirms a general upward trend in geographic concentration of industries with slow formation of clusters around several locations.

The situation is slightly better in two industrial groups I devote my particular attention to, in machinery and high tech. The number of raions with only one plant in the group is about forty percent. Activity in both of these groups follows a general trend in manufacturing and tends to relocate to bigger cities: the total number of establishments increased over five years, whereas the number of raions where an activity is located has dropped.

A small number of plants in machinery and high tech (between 1.5% in 2001 and 2.1% in 2005) are foreign owned. However, there are differences in the shares of workers employed by foreign-owned firms. In the high-tech, the employment share is also quite small and does not exceed three percent of total employment in the group. In machinery, the share of employment at foreign-owned firms has been increasing steadily from 2.5% in 2001 to 9.6% in 2005. This implies that foreigner owners are more interested in big machinery firms. I cannot say so far if the spatial concentration foreign ownership will have any implications for localization economies. Kyiv as the capital of the country attracts most such firms, and still their concentration is quite thin. In the rest of the country the concentration of foreign ownership is even sparser. However if I find empirical support to the fact that foreign ownership is able to benefit more from the agglomeration economies, I may expect an increase of foreign-owned firms around several cities and, as a result, general increase in the localization effects for all firms in such locations.

My analysis of agglomeration indices also reveals a slow increase in concentration both in employment the number of plants. I also found certain resemblance of agglomeration patterns in Ukraine with those in Italy. This leads to a conclusion that Ukrainian economy is starting to reveal features common in market economies in more developed countries.

In my analysis I left aside several possible extensions for the future research. First, the composition of the lattice data in Ukraine (and, as a result, the spatial weight matrix) is quite unusual due to a significant number of perforated regions and cities having the same spatial status as larger rural raions. Such spatial structure leaves too many regions with only a single neighbor and increases inequality in the levels of economic activity between two neighboring regions (such as a city and a surrounding rural raion). This particular case of the Modifiable Areal Unit Problem demands special treatment. Partitioning the space into more equal entities based, for example, on commuting patterns may provide further insight into the nature of spatial correlation of economic activity.

Second, all existing measures of geographic concentration, such as Location Quotients or Ellison-Glaeser type indices, treat all regions as spatially independent. The value of the index does not depend on a particular form of the adjacency matrix for the economy. One of interesting experiments would be to combine smaller regions into economically sound clusters first, and then perform the analysis of geographic concentration with clustered data.

Geographic concentration of manufacturing is apparently increasing in Ukraine over time. The major issue is how to measure this increase and clear it from confounding factors.

Tables

Table 3.1, Dynamics of the Total Employment (Thousands) by Regions, 1990-2005

	Year						
Region	1990	2000	2001	2003	2005	1990 to 2005 Change	2000 to 2005 Change
Center	5,435	3,210	3,008	2,620	2,461	-55%	-23%
East	8,670	5,121	4,886	4,531	4,427	-49%	-14%
South	3,417	1,912	1,784	1,617	1,540	-55%	-19%
West	4,470	2,591	2,397	2,077	1,995	-55%	-23%
Kyiv*	1,375	846	856	866	965	-30%	14%
Ukraine	23,367	13,680	12,931	11,711	11,388	-51%	-17%

Source: Derzhkomstat (2006), Table 17.27

Table 3.2, Dynamics of the Total Employment by Sectors, 2000-2005

	2000	2001	2003	2005	2000 – 2005 change
Total	13,678	12,931	11,711	11,388	-16.7%
Agriculture, Forestry	2,551	2,206	1,537	1,137	-55.4%
Fishing	31	28	23	16	-48.4%
Manufacturing	4,061	3,811	3,416	3,416	-15.9%
Construction	590	526	431	461	-21.9%
Wholesale and Retail Trade, Repair	630	600	563	676	7.3%
Hotels and Restaurants	100	91	81	82	-18.0%
Transportation and communication	1,110	1,048	994	992	-10.6%
Financial Activity	146	151	169	216	47.9%
Real Estate and Mortgages	596	594	560	564	-5.4%
Public Administration	610	631	702	570	-6.6%
Education	1,551	1,565	1,584	1,610	3.8%
Health Care and Social Services	1,304	1,291	1,281	1,271	-2.5%
Personal Sericves	398	390	370	377	-5.3%

Source: Derzhkomstat (2006), Table 17.26

Table 3.3, Dynamics of Manufacturing Employment by Subsectors, 2000-2005.

	Years						
	2000	2001	2002	2003	2004	2005	2000 - 2005 change
DA: Food, beverages and tobacco	518	485	464	445	452	465	-10%
DB: Textiles and products	221	197	161	135	129	126	-43%
DC: Leather and products	41	33	25	22	22	21	-49%
DD: Wood and wood products	41	36	34	33	34	37	-10%
DE: Pulp, paper, publishing	60	61	58	57	61	66	10%
DF: Coke, refined petroleum products	54	55	57	58	59	59	9%
DG: Chemicals, man-made fibres	159	153	152	140	135	134	-16%
DH: Rubber and plastic	42	40	39	39	43	48	14%
DI: Other non-metallic mineral product	203	184	161	150	150	154	-24%
DJ: Basic metals and fabricated metal	434	428	418	409	415	425	-2%
DK: Machinery and equipment n.e.c.	497	462	415	378	361	352	-29%
DL: Electrical and optical equipment	281	239	205	184	180	174	-38%
DM: Transport equipment	283	257	247	247	252	247	-13%
DN: Manufacturing n.e.c.	83	74	64	59	60	63	-24%
Total Manufacturing	2917	2704	2500	2356	2351	2370	-19%

Source: Derzhkomstat (2006), Table 17.28.

Table 3.4, Shares of Urban Employment and Plant Counts for Entire Economy

Year	Employment			Manufacturing Employment		
	Urban	Total		Urban	Total	
2001	6,789,746	10,297,236	66%	2,443,166	3,040,958	80%
2002	6,720,397	9,926,698	68%	2,314,524	2,856,757	81%
2003	6,662,454	9,499,778	70%	2,214,158	2,719,937	81%
2004	6,789,134	9,346,070	73%	2,218,478	2,710,702	82%
2005	6,856,188	9,213,782	74%	2,177,751	2,670,878	82%
	Plant Count			Manufacturing Plant Count		
	Urban	Total		Urban	Total	
2001	165,710	222,262	75%	26,190	35,989	73%
2002	180,820	239,800	75%	27,927	38,040	73%
2003	190,917	251,605	76%	28,739	39,076	74%
2004	196,326	256,146	77%	28,880	38,780	74%
2005	201,634	260,483	77%	29,115	38,634	75%

Table 3.5, Moran's I Test of Spatial Correlation in Manufacturing

	I(adj)	z(adj)	I(adj)*	z(adj)*
Employment				
2001	0.024	0.976	0.025	0.992
2002	0.026	1.065	0.026	1.053
2003	0.029	1.189	0.030	1.191
2004	0.031	1.266	0.031	1.231
2005	0.036	1.451	0.037	1.467
Plants count				
2001	0.029	1.380	0.032	1.317
2002	0.032	1.519	0.035	1.427
2003	0.034	1.601	0.036	1.462
2004	0.034	1.649	0.037	1.526
2005	0.035	1.764	0.038	1.533

Table 3.6, Average Size of Plants

Year	2001	2002	2003	2004	2005
DA: Food, beverages and tobacco	75.6	71.6	67.5	73.4	79.3
DB: Textiles and products	69.7	62.2	49.8	50.3	50.6
DC: Leather and products	85.3	64.7	53.1	64.7	64.8
DD: Wood and wood products	19.9	18.2	18.2	17.9	18.4
DE: Pulp, paper, publishing	20.2	18.7	17.8	16.3	16.5
DF: Coke, refined petroleum products	443	216.2	189.2	167.5	224.8
DG: Chemicals, man-made fibres	144.7	142.9	130.8	127.4	122.3
DH: Rubber and plastic	39.5	40.2	36.6	35.9	38.5
DI: Other non-metallic mineral product	93.6	76.2	66.5	66.1	63.6
DJ: Basic metals and fabricated metal	117.4	104.1	93.9	99.5	93.4
DK: Machinery and equipment n.e.c.	94.4	82.9	74.3	70.6	74.9
DL: Electrical and optical equipment	48.7	42.2	41.1	37.4	45.1
DM: Transport equipment	268.9	230.2	201.7	238.9	220.4
DN: Manufacturing n.e.c.	47.6	40.4	36.4	36.3	36.1

Table 3.7, Number of Raions with Plants in a Subsector

Year	2001	2002	2003	2004	2005
DA: Food, beverages and tobacco	655	658	656	652	648
DB: Textiles and products	463	439	424	406	405
DC: Leather and products	123	113	116	105	106
DD: Wood and wood products	384	406	424	432	423
DE: Pulp, paper, publishing	562	575	577	576	579
DF: Coke, refined petroleum products	68	139	147	172	128
DG: Chemicals, man-made fibres	217	226	220	216	223
DH: Rubber and plastic	202	202	209	210	221
DI: Other non-metallic mineral product	409	435	454	456	467
DJ: Basic metals and fabricated metal	308	324	342	341	337
DK: Machinery and equipment n.e.c.	531	522	508	484	470
DL: Electrical and optical equipment	404	399	390	382	366
DM: Transport equipment	154	166	162	159	155
DN: Manufacturing n.e.c.	344	335	321	317	323

Table 3.8, Number of Raions with Location Quotient (Employment-Based) Above and Below One

	LQ ≤1					LQ >1				
Year	2001	2002	2003	2004	2005	2001	2002	2003	2004	2005
DA: Food, beverages and tobacco	169	172	173	186	177	486	486	483	466	471
DB: Textiles and products	277	244	240	226	225	186	195	184	180	180
DC: Leather and products	59	50	58	47	49	64	63	58	58	57
DD: Wood and wood products	210	212	221	221	218	174	194	203	211	205
DE: Pulp, paper, publishing	362	369	369	369	386	201	206	208	207	193
DF: Coke, refined petroleum products	41	98	95	119	86	27	41	52	53	42
DG: Chemicals, man-made fibres	155	168	156	152	152	62	58	64	64	72
DH: Rubber and plastic	123	111	117	117	134	79	91	92	93	87
DI: Other non-metallic mineral product	235	254	259	258	266	174	181	195	198	201
DJ: Basic metals and fabricated metal	254	272	279	276	275	54	52	63	65	62
DK: Machinery and equipment n.e.c.	333	347	341	328	321	198	175	167	156	149
DL: Electrical and optical equipment	298	307	294	286	277	106	92	97	96	89
DM: Transport equipment	103	110	106	103	100	51	56	56	56	55
DN: Manufacturing n.e.c.	208	190	192	194	197	136	145	129	123	126

Table 3.9, Number of Raions with Location Quotient (Plant-Based) Above and Below One

	LQ ≤1					LQ >1				
Year	2001	2002	2003	2004	2005	2001	2002	2003	2004	2005
DA: Food, beverages and tobacco	132	138	136	136	132	523	520	520	516	516
DB: Textiles and products	204	193	196	172	166	259	246	228	234	239
DC: Leather and products	25	29	28	25	22	98	84	88	80	84
DD: Wood and wood products	155	164	174	148	139	229	242	250	284	284
DE: Pulp, paper, publishing	328	366	357	336	346	235	209	220	240	233
DF: Coke, refined petroleum products	9	14	12	16	14	59	125	135	156	114
DG: Chemicals, man-made fibres	54	71	71	72	69	163	155	149	144	155
DH: Rubber and plastic	58	60	75	69	86	144	142	134	141	135
DI: Other non-metallic mineral product	136	141	125	140	151	273	294	329	316	316
DJ: Basic metals and fabricated metal	149	159	172	179	161	159	165	170	162	176
DK: Machinery and equipment n.e.c.	248	261	257	254	227	283	261	251	230	243
DL: Electrical and optical equipment	249	247	237	250	236	155	152	154	132	130
DM: Transport equipment	49	52	48	51	50	105	114	114	108	105
DN: Manufacturing n.e.c.	159	159	152	159	149	185	176	169	158	174

Table 3.10, Descriptive Statistics of Machinery and High Tech Groups

Year	Number of raions with activity	Total number of plants within groups	Total employment within groups	Average number of plants per raion	Average employment per firm
			Machinery		
2001	361	3,117	343,930	8.6	110.3
2002	366	3,297	339,177	9	102.9
2003	352	3,467	309,416	9.8	89.2
2004	340	3,448	291,956	10.1	84.7
2005	331	3,363	296,841	10.2	88.3
			High Tech		
2001	160	1,037	138,065	6.5	133.1
2002	158	1,098	103,536	6.9	94.3
2003	152	1,044	92,070	6.8	88.2
2004	141	1,049	94,194	7.4	89.8
2005	144	1,042	77,622	7.2	74.5

Table 3.11, Moran's I Test for Global Spatial Autocorrelation

Year	Machinery		High Tech	
	I	z(I)	I	z(I)
	Employment			
2001	0.035	1.389	0.006	0.321
2002	0.030	1.251	0.017	0.777
2003	0.015	0.668	0.012	0.571
2004	0.017	0.725	0.011	0.526
2005	0.018	0.748	-0.006	-0.183
	Number of plants			
2001	0.022	0.964	0.023	1.294
2002	0.019	0.854	0.017	1.104
2003	0.020	0.892	0.013	0.875
2004	0.016	0.718	0.003	0.312
2005	0.014	0.641	0.005	0.425

Table 3.12, Ownership Structure of Machinery and High Tech Groups

Year	Number of plants	Private		Foreign Owned	
Machinery					
2001	3,117	2,934	(94.1%)	41	(1.3%)
2002	3,297	3,121	(94.7%)	47	(1.4%)
2003	3,467	3,304	(95.3%)	51	(1.5%)
2004	3,448	3,291	(95.4%)	55	(1.6%)
2005	3,363	3,228	(96.0%)	68	(2.0%)
High Tech					
2001	1,037	947	(91.3%)	16	(1.5%)
2002	1,098	1,011	(92.1%)	21	(1.9%)
2003	1,044	961	(92.0%)	19	(1.8%)
2004	1,049	965	(92.0%)	24	(2.3%)
2005	1,042	957	(91.8%)	21	(2.0%)

Year	Employment	Private		Foreign Owned	
Machinery					
2001	343,930	276,598	(80.4%)	8,184	(2.4%)
2002	339,177	264,953	(78.1%)	23,586	(7.0%)
2003	309,416	247,988	(80.1%)	25,357	(8.2%)
2004	291,956	246,455	(84.4%)	26,269	(9.0%)
2005	296,841	248,948	(83.9%)	28,433	(9.6%)
High Tech					
2001	138,065	58,524	(42.4%)	768	(0.6%)
2002	103,536	53,672	(51.8%)	1,225	(1.2%)
2003	92,070	46,397	(50.4%)	1,171	(1.3%)
2004	94,194	42,781	(45.4%)	1,227	(1.3%)
2005	77,622	44,287	(57.1%)	1,772	(2.3%)

Table 3.13, Number of Raions with Private and Foreign Firms in Machinery and High Tech

Year	Total	Private	Foreign
Machinery			
2001	361	350	21
2002	366	358	25
2003	352	341	34
2004	340	330	32
2005	331	323	42
High Tech			
2001	160	156	9
2002	158	151	9
2003	153	145	10
2004	141	135	12
2005	144	139	11

Table 3.14, Dynamics of Employment –Based (Weighted) Ellison-Glaeser Indices

Year	Mean	Median	Std. dev.	Negative	Small	Medium	Large
2001	0.0163	0.0059	0.0419	28	50	12	13
2002	0.0213	0.0058	0.0426	23	50	14	16
2003	0.0194	0.0074	0.0445	28	46	14	15
2004	0.0209	0.0069	0.0606	26	51	14	12
2005	0.0230	0.0082	0.0653	22	50	17	14

Table 3.15, Employment-Based Concentration by Manufacturing Subsectors, 2001 and 2005

		2001					2005				
Subsector	Total	N	S	M	L	Size	N	S	M	L	Size
DA: Food, beverages and tobacco	10	2	8	0	0	80	2	8	0	0	92
DB*: Textiles and Leather	13	4	7	2	0	117	3	7	2	1	73
DD: Wood and wood products	5	0	3	2	0	46	1	3	1	0	47
DE: Pulp, paper, publishing	5	0	2	1	2	81	0	1	1	3	64
DG*: Chemicals, fibres, coke, ref. pertol.	10	4	4	1	1	368	2	5	2	1	289
DH: Rubber and plastic	2	0	1	0	1	90	0	1	0	1	72
DI: Other non-metallic mineral product	8	2	3	2	1	131	1	6	0	1	107
DJ: Basic metals and fabricated metal	12	2	8	0	2	828	3	6	3	0	357
DK: Machinery and equipment n.e.c.	7	1	6	0	0	283	3	3	0	1	136
DL: Electrical and optical equipment	15	5	4	2	4	81	2	5	5	3	55
DM: Transport equipment	8	6	1	0	1	308	4	2	0	2	267
DN: Manufacturing n.e.c.	8	2	3	2	1	42	1	3	3	1	36
Total	103	28	50	12	13	231	22	50	17	14	144

Table 3.16, Dynamics of Plant –Based (Unweighted) Ellison-Glaeser Indices

Year	Mean	Median	Std. dev.	Negative	Small	Medium	Large
2001	0.0171	0.0087	0.0367	8	74	11	10
2002	0.0211	0.0080	0.0496	10	71	11	11
2003	0.0173	0.0097	0.0253	7	72	16	8
2004	0.0178	0.0082	0.0326	8	75	10	10
2005	0.0193	0.0128	0.0365	10	69	15	9

Table 3.17, Ten Most and Least Concentrated Industries in 2001 and 2005, Weighted Gamma.

KVED		Gamma	Size	Plants	Regions
	2001				
221	Publishing	0.2108	21	1721	381
251	Rubber products	0.1519	153	169	63
351	Building and repairing of ships and boats	0.1247	189	257	23
300	Office machinery and computers	0.1218	24	445	103
334	Optical instruments and photographic equipment	0.1199	159	50	23
314	Accumulators, primary cells and primary batteries	0.1186	98	39	22
272	Manufacture of tubes	0.1162	1,185	40	21
223	Reproduction of recorded media	0.0914	12	44	22
363	Manufacture of musical instruments	0.0802	34	23	12
267	Cutting, shaping and finishing of stone	0.0636	20	291	104
	2005				
223	Reproduction of recorded media	0.3439	20	31	11
363	Manufacture of musical instruments	0.3156	44	11	9
335	Manufacture of watches and clocks	0.2593	14	6	3
183	Dressing and dyeing of fur; articles of fur	0.1953	42	50	25
300	Manufacture of office machinery and computers	0.1938	22	384	85
221	Publishing	0.1273	18	2,343	487
351	Building and repairing of ships and boats	0.1025	120	312	25
334	Optical instruments and photographic equipment	0.1023	111	47	22
222	Printing and service activities related to printing	0.0658	16	2,564	482
244	Pharmaceuticals, med. chemicals.	0.0646	86	254	75

KVED		Gamma	Size	Plants	Regions
	2001				
233	Processing of nuclear fuel	-0.0923	220	6	6
173	Finishing of textiles	-0.0850	275	7	6
354	Manufacture of motorcycles and bicycles	-0.0404	118	26	9
160	Manufacture of tobacco products	-0.0355	200	30	25
335	Manufacture of watches and clocks	-0.0284	13	7	6
341	Manufacture of motor vehicles	-0.0214	198	194	81
353	Manufacture of aircraft and spacecraft	-0.0204	1,119	63	21
172	Textile weaving	-0.0179	383	53	31
263	Manufacture of ceramic tiles and flags	-0.0142	94	53	32
268	Other non-metallic mineral products	-0.0114	87	179	85
	2005				
354	Manufacture of motorcycles and bicycles	-0.1777	563	20	10
173	Finishing of textiles	-0.0745	21	7	7
263	Manufacture of ceramic tiles and flags	-0.0621	112	42	27
353	Manufacture of aircraft and spacecraft	-0.0392	517	78	24
233	Processing of nuclear fuel	-0.0390	71	7	6
341	Manufacture of motor vehicles	-0.0355	150	175	70
160	Manufacture of tobacco products	-0.0325	225	19	15
275	Casting of metals	-0.0223	506	166	77
342	Bodies for motor vehicles; trailers	-0.0158	88	44	29
172	Textile weaving	-0.0122	179	55	27

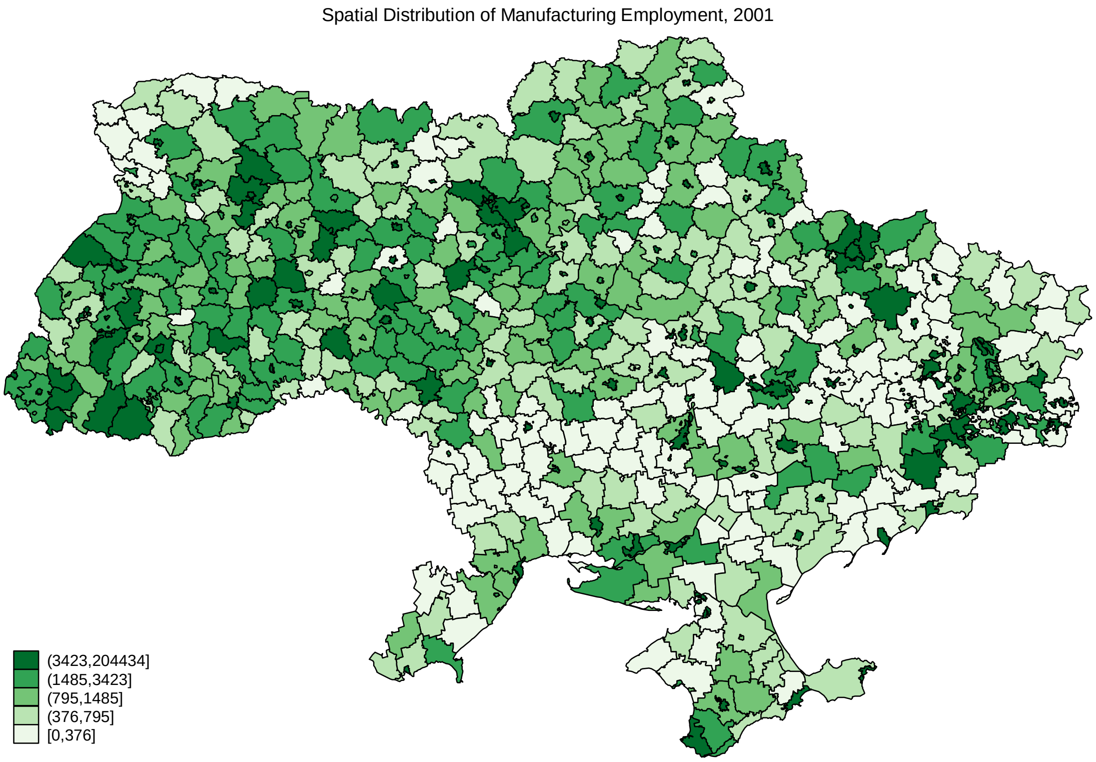
Table 3.18, Ten Most and Least Concentrated Industries in 2001 and 2005, Unweighted Gamma.

KVED		Gamma	Size	Plants	Regions
	2001				
176	Manufacture of knitted and crocheted fabrics	0.2733	85	50	19
351	Building and repairing of ships and boats	0.1475	189	257	23
353	Manufacture of aircraft and spacecraft	0.1140	1,119	63	21
354	Manufacture of motorcycles and bicycles	0.1000	118	26	9
313	Manufacture of insulated wire and cable	0.0747	107	87	27
364	Manufacture of sports goods	0.0694	34	34	13
231	Manufacture of coke oven products	0.0684	1,474	22	12
267	Cutting, shaping and finishing of stone	0.0555	20	291	104
247	Manufacture of man-made fibres	0.0548	1,136	10	6
272	Manufacture of tubes	0.0545	1,185	40	21
	2005				
223	Reproduction of recorded media	0.1795	20	31	11
335	Manufacture of watches and clocks	0.1792	14	6	3
353	Manufacture of aircraft and spacecraft	0.1755	517	78	24
351	Building and repairing of ships and boats	0.1622	120	312	25
272	Manufacture of tubes	0.0811	1,217	45	22
247	Manufacture of man-made fibres	0.0763	532	11	8
364	Manufacture of sports goods	0.0703	28	50	17
313	Insulated wire and cable	0.0600	95	89	32
334	Optical instruments and photographic equip.	0.0579	111	47	22
221	Publishing	0.0459	18	2343	487

KVED		Gamma	Size	Plants	Regions
	2001				
233	Processing of nuclear fuel	-0.0597	220	6	6
335	Manufacture of watches and clocks	-0.0452	13	7	6
173	Finishing of textiles	-0.0357	275	7	6
372	Recycling of non-metal waste and scrap	-0.0029	23	150	78
160	Manufacture of tobacco products	-0.0028	200	30	25
263	Manufacture of ceramic tiles and flags	-0.0020	94	53	32
172	Textile weaving	-0.0011	383	53	31
243	Paints, varnishes coatings, ink and mastics	-0.0004	40	226	82
175	Manufacture of other textiles	0.0001	82	102	52
203	Manufacture of builders carpentry and joinery	0.0003	16	1395	315
	2005				
173	Finishing of textiles	-0.0644	21	7	7
355	Manufacture of other transport equipment n.e.c.	-0.0182	19	14	11
176	Manufacture of knitted and crocheted fabrics	-0.0164	150	16	13
296	Manufacture of weapons and ammunition	-0.0120	379	18	12
233	Processing of nuclear fuel	-0.0102	71	7	6
342	Bodies for motor vehicles; trailers	-0.0062	88	44	29
263	Manufacture of ceramic tiles and flags	-0.0037	112	42	27
175	Manufacture of other textiles	-0.0023	50	108	57
211	Manufacture of pulp, paper and paperboard	-0.0005	214	46	34
287	Manufacture of other fabricated metal products	0.0000	48	723	186

Figures

Figure 3.1, Change in Total Employment between 2000 and 2005



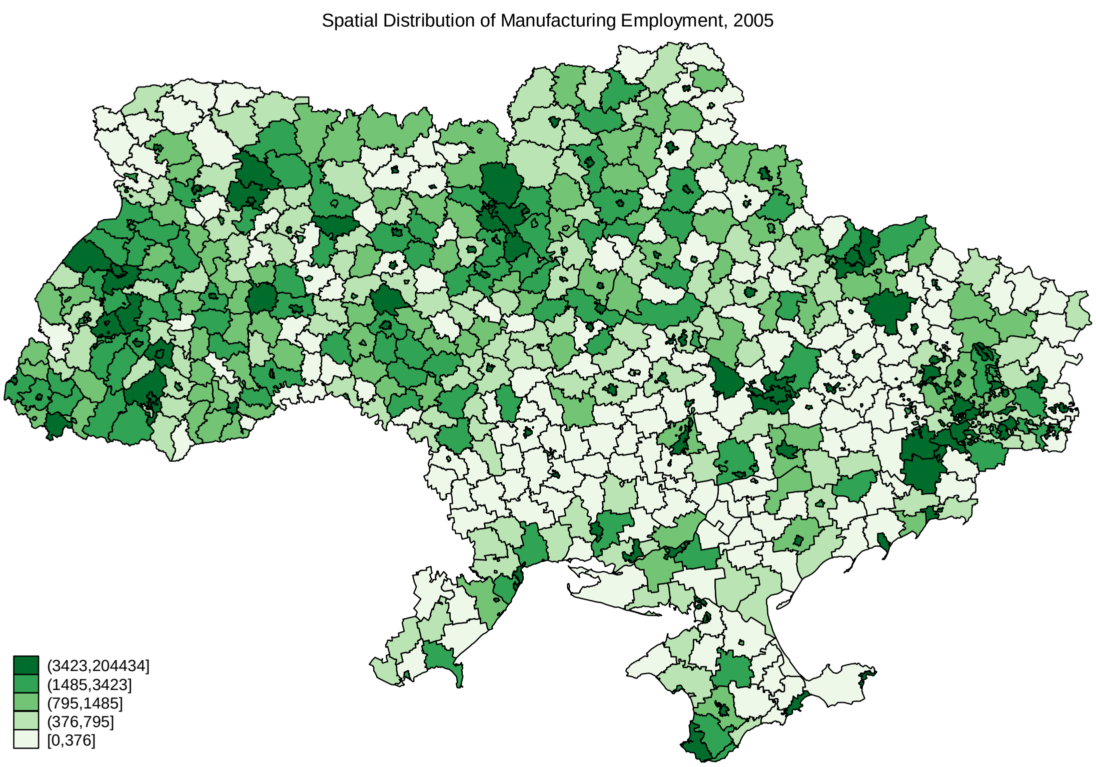


Figure 3.2, Manufacturing Employment in Ukrainian Raions, 2001 and 2005.

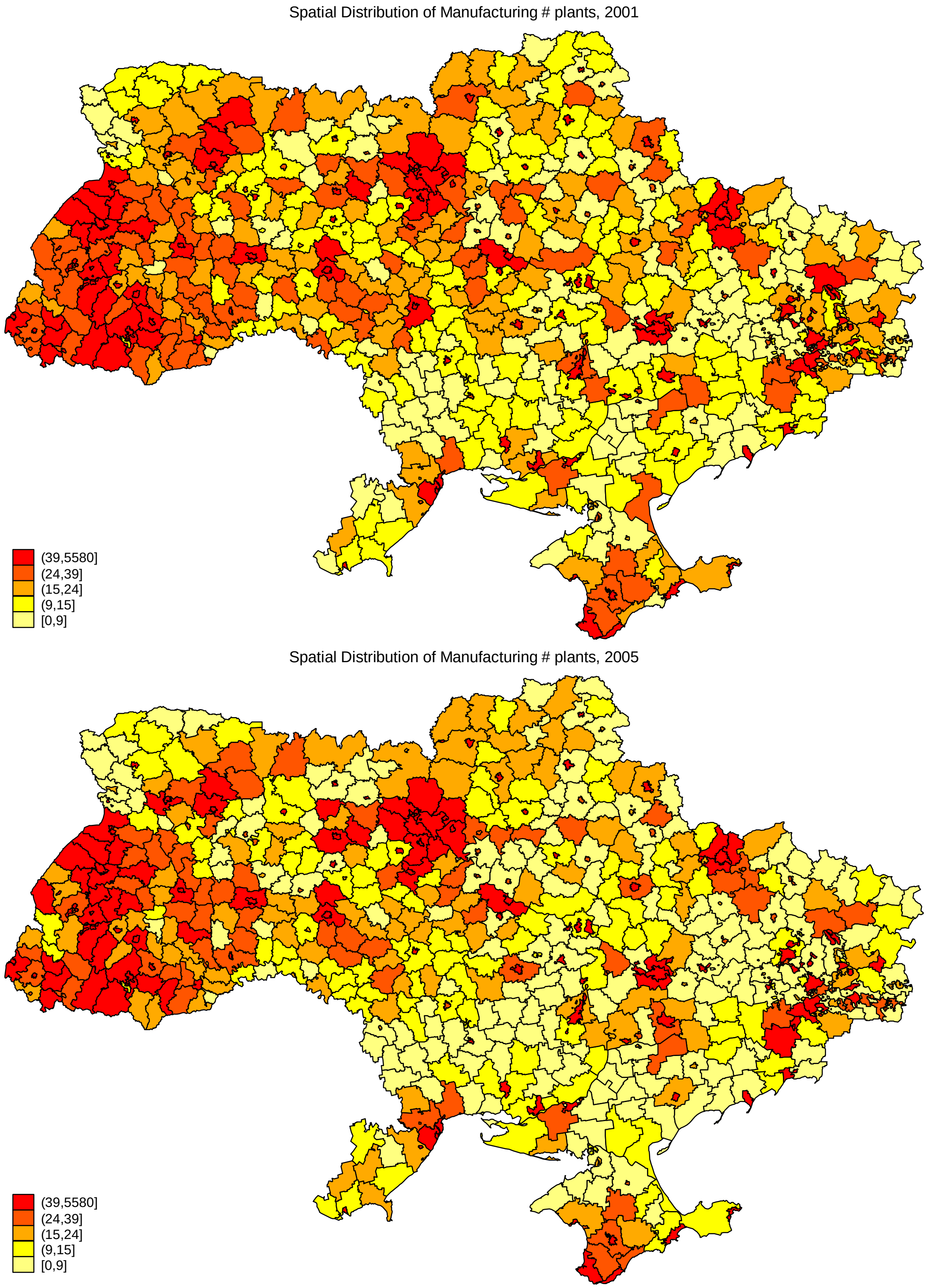


Figure 3.3, Manufacturing Plants in Ukrainian Raions, 2001 and 2005.

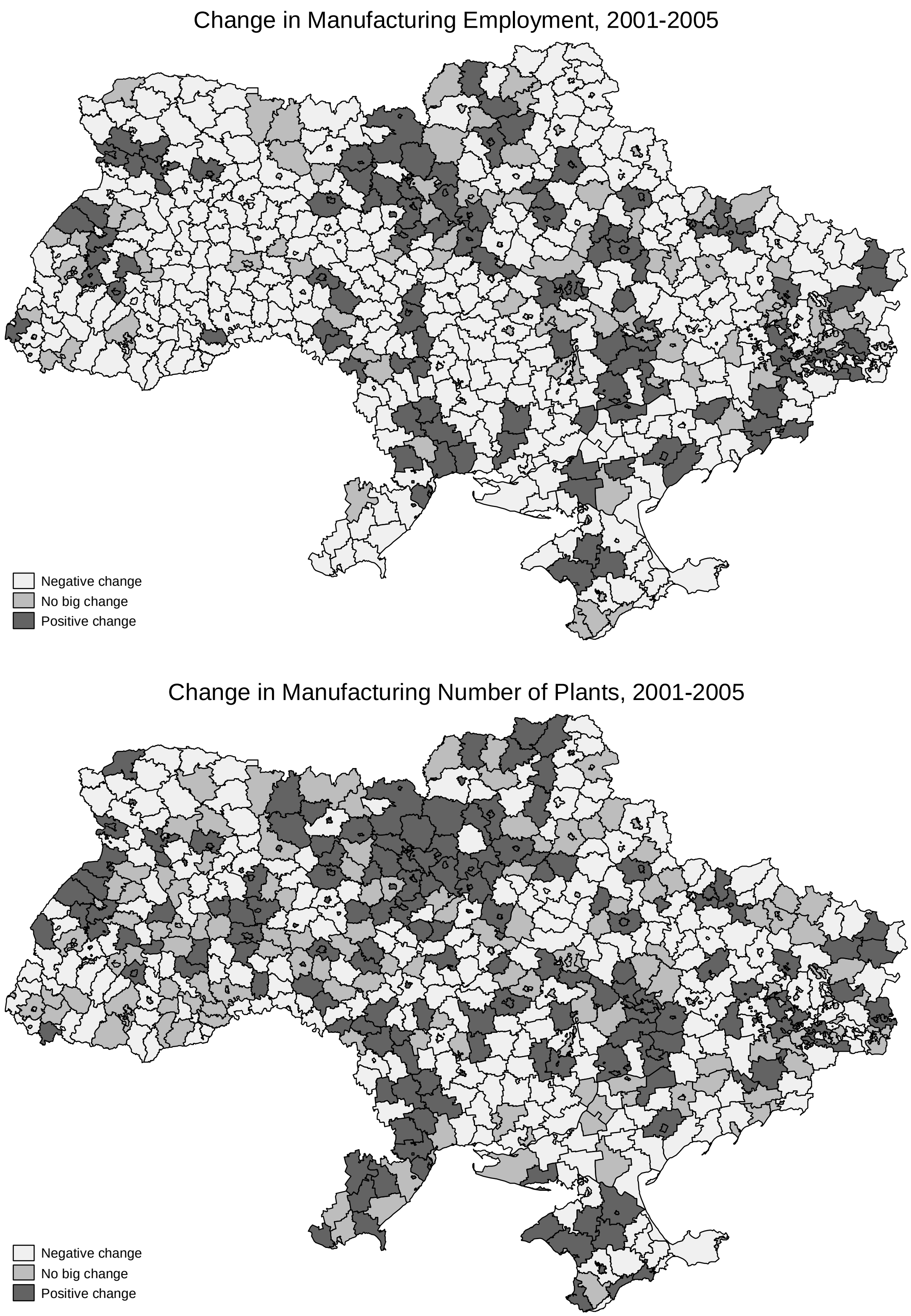


Figure 3.4, Changes in the Manufacturing Employment and the Number of Plants, 2001 – 2005

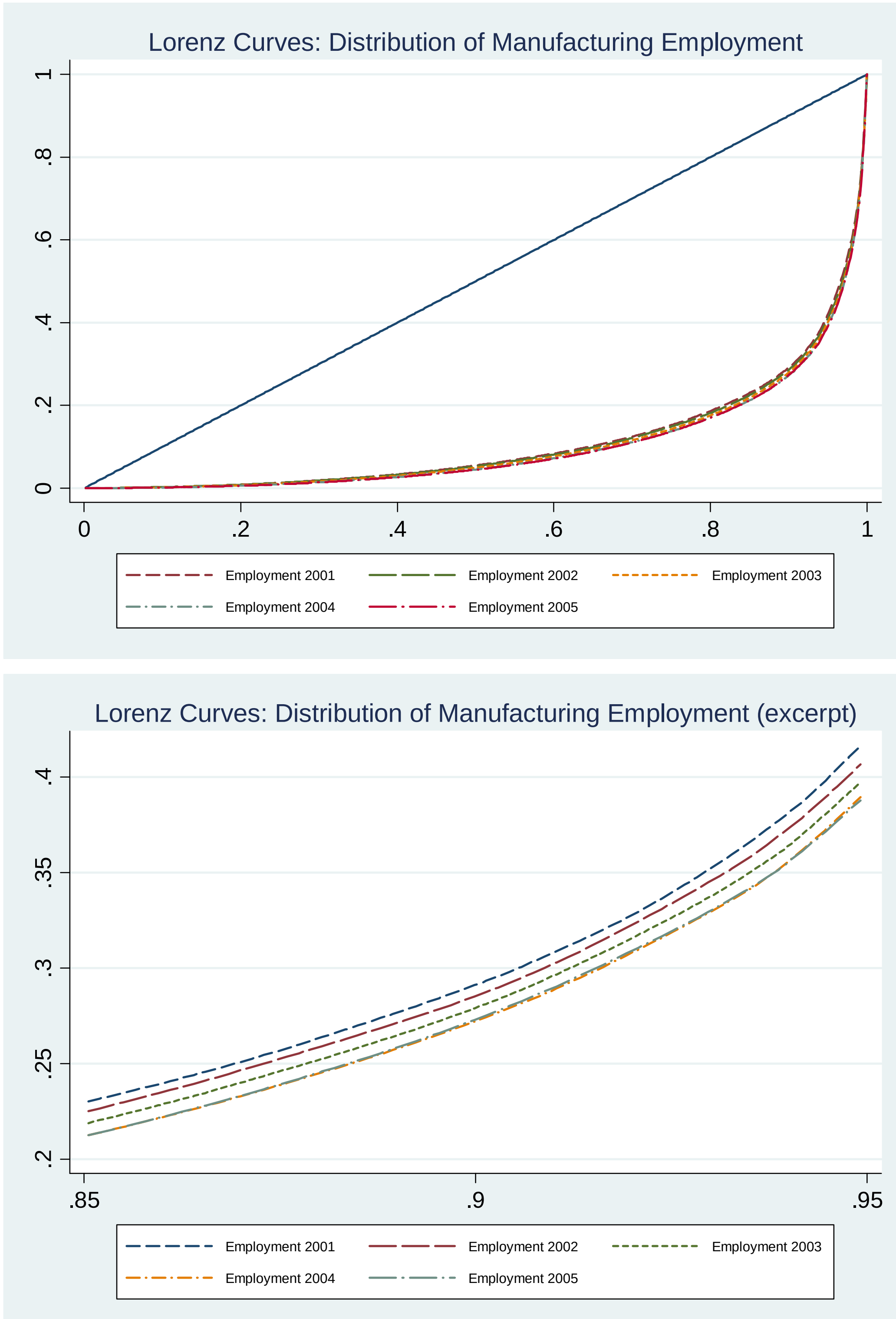


Figure 3.5, Lorenz Curves: Distribution of Manufacturing Employment Across Raions

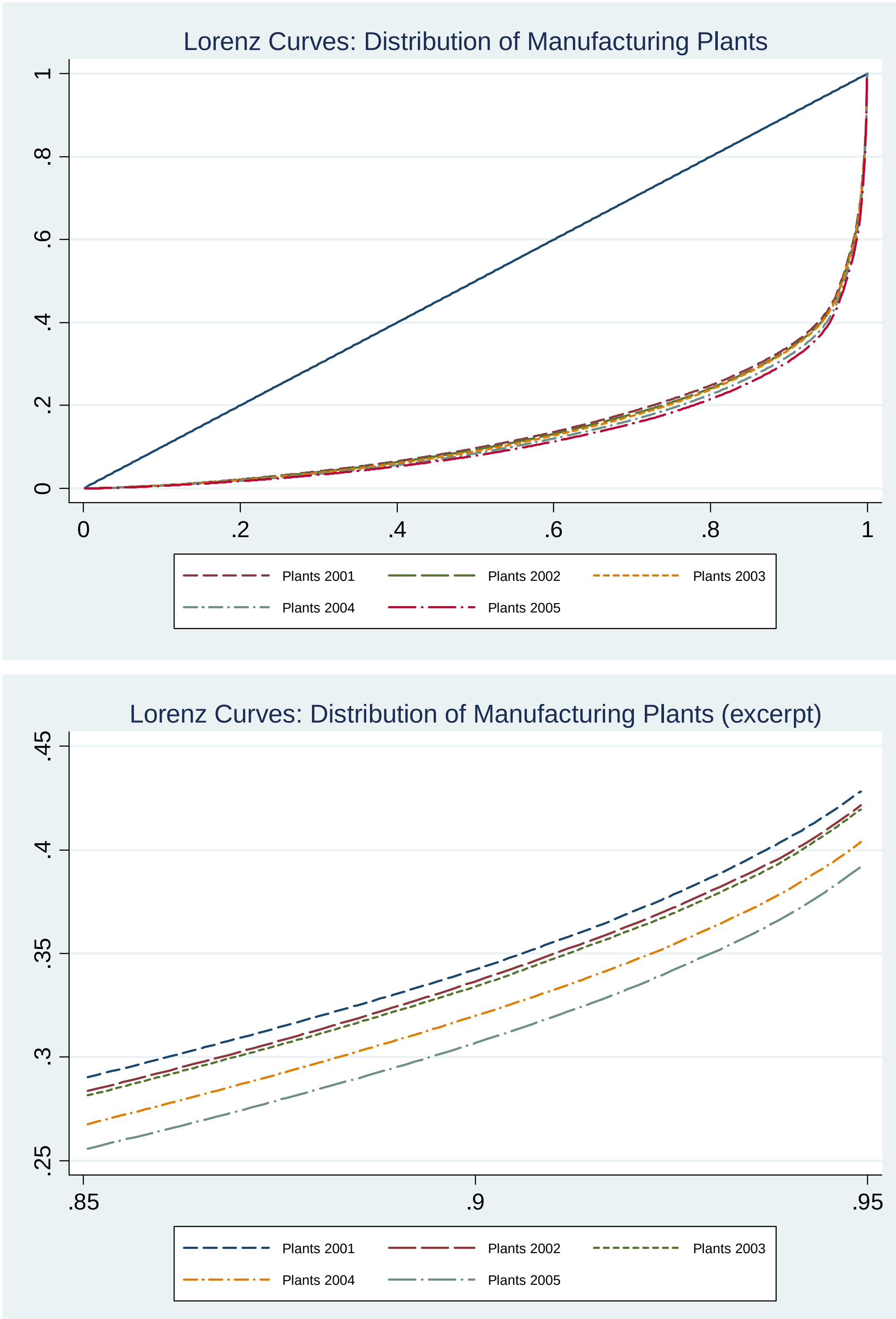


Figure 3.6, Lorenz Curves: Distribution of Manufacturing Plants across Raions

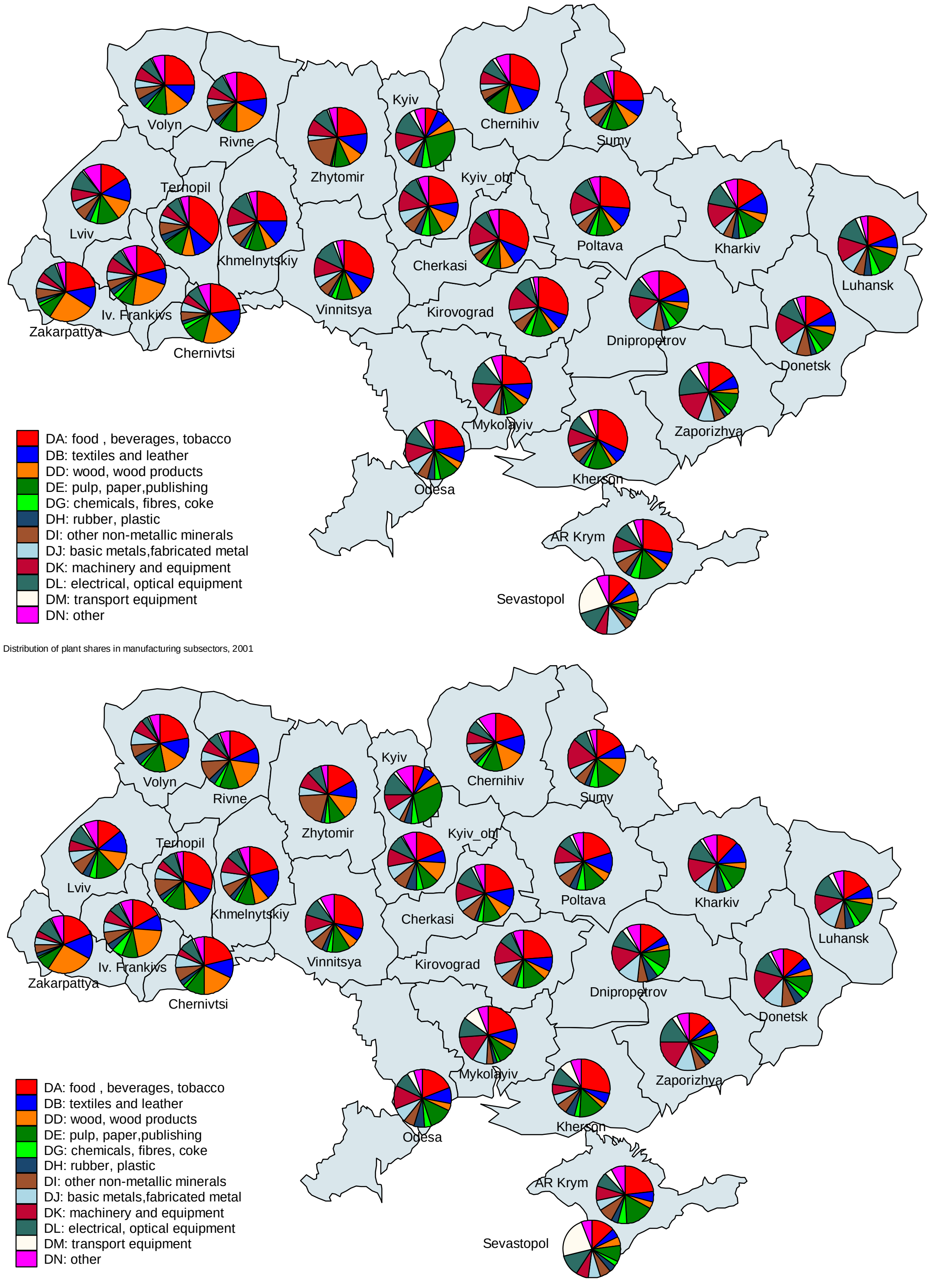


Figure 3.7, Distribution of Plant Shares in Manufacturing Subsectors, 2001 and 2005

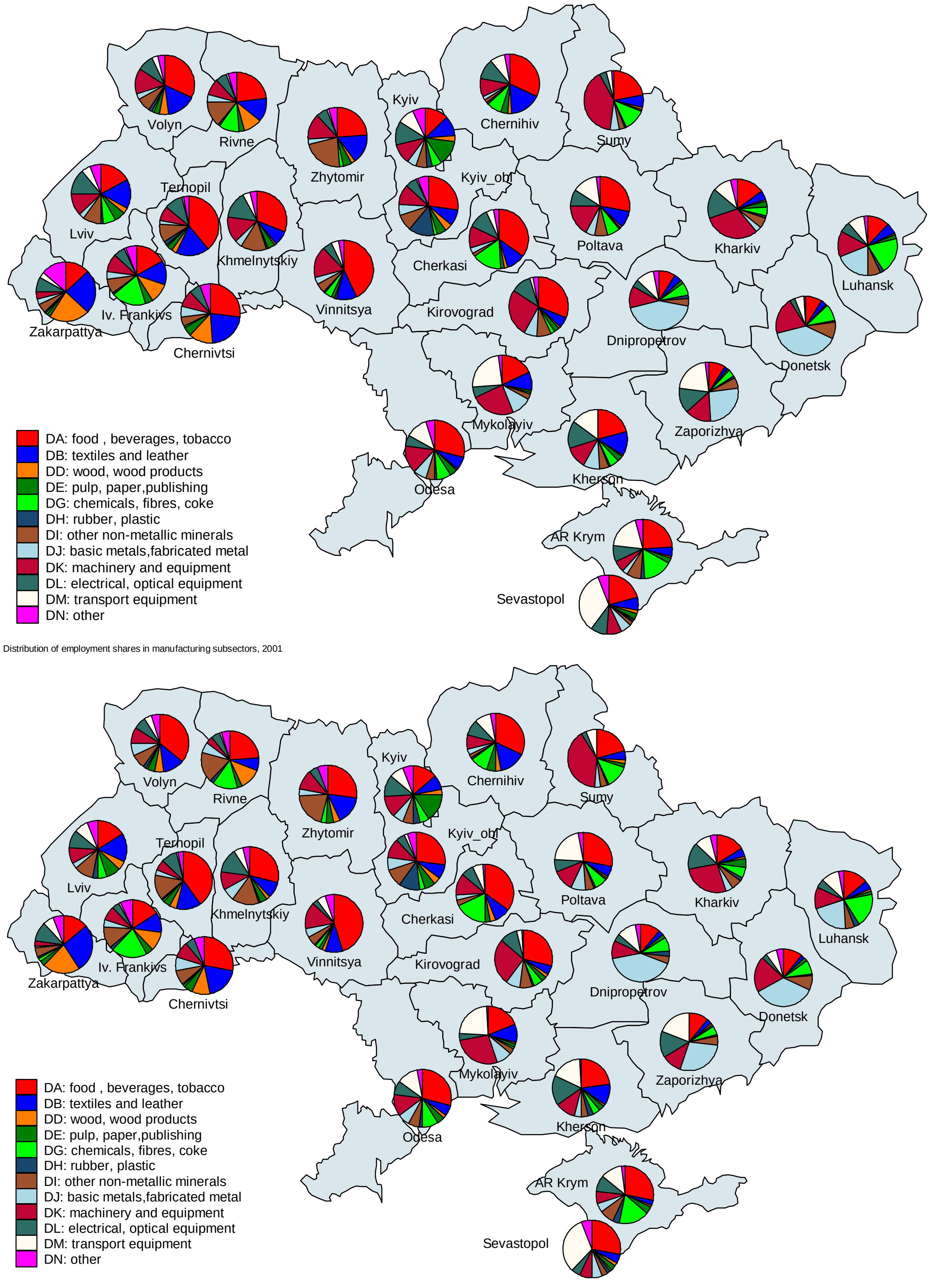
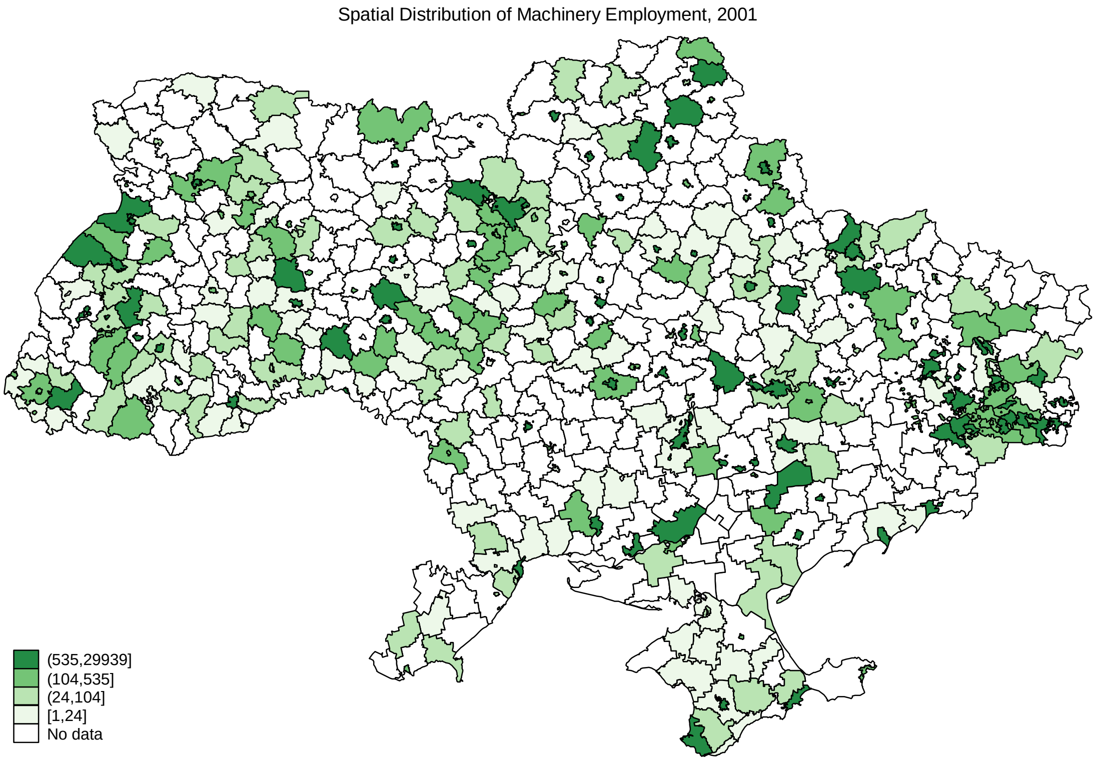


Figure 3.8, Distribution of Employment Shares in Manufacturing Subsectors, 2001 and 2004



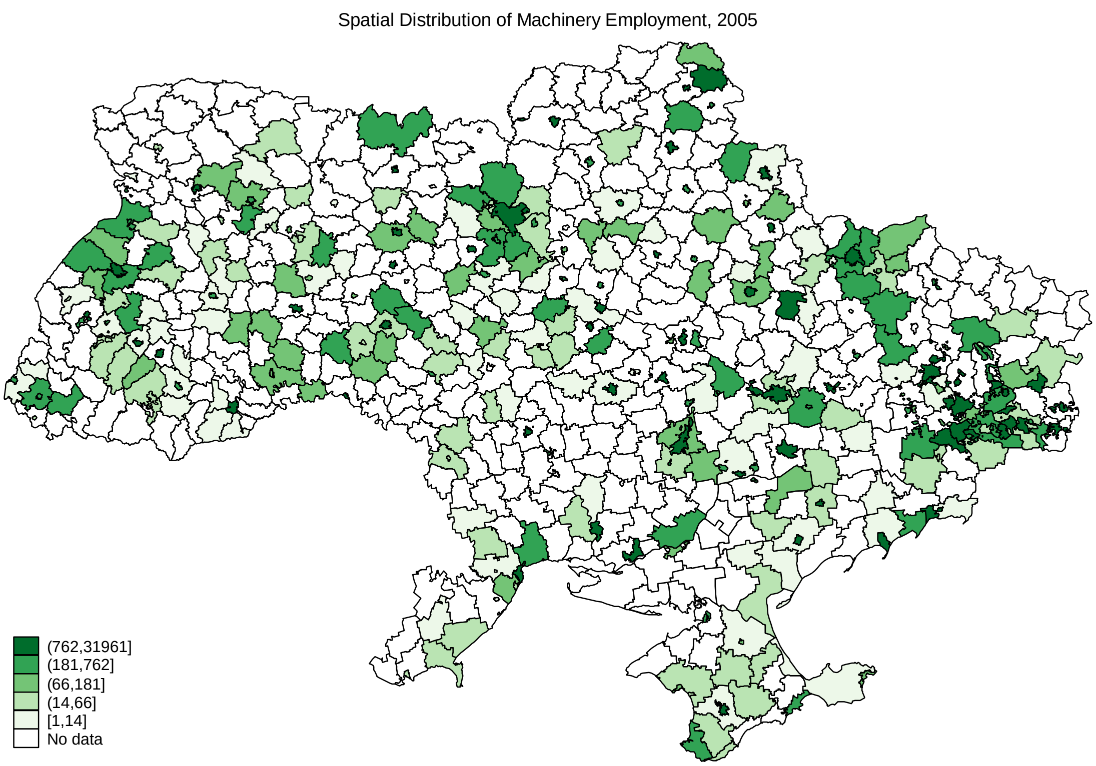
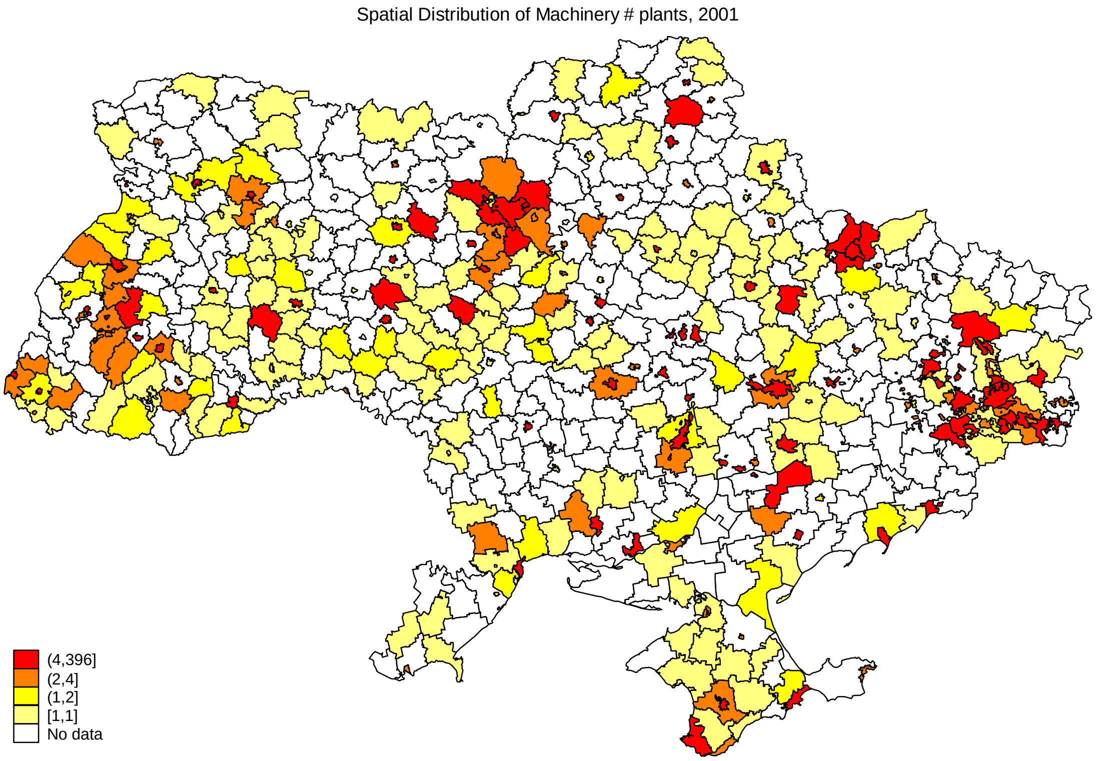


Figure 3.9, Spatial Distribution of Employment in Machinery, 2001 and 2005



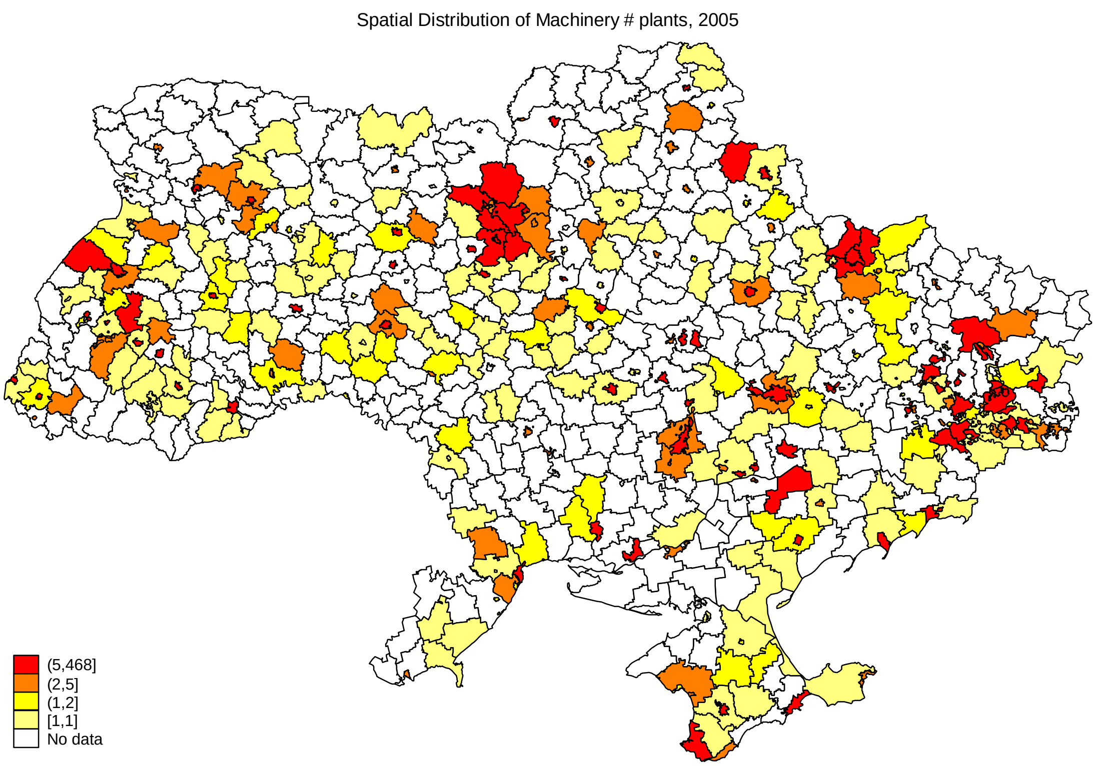
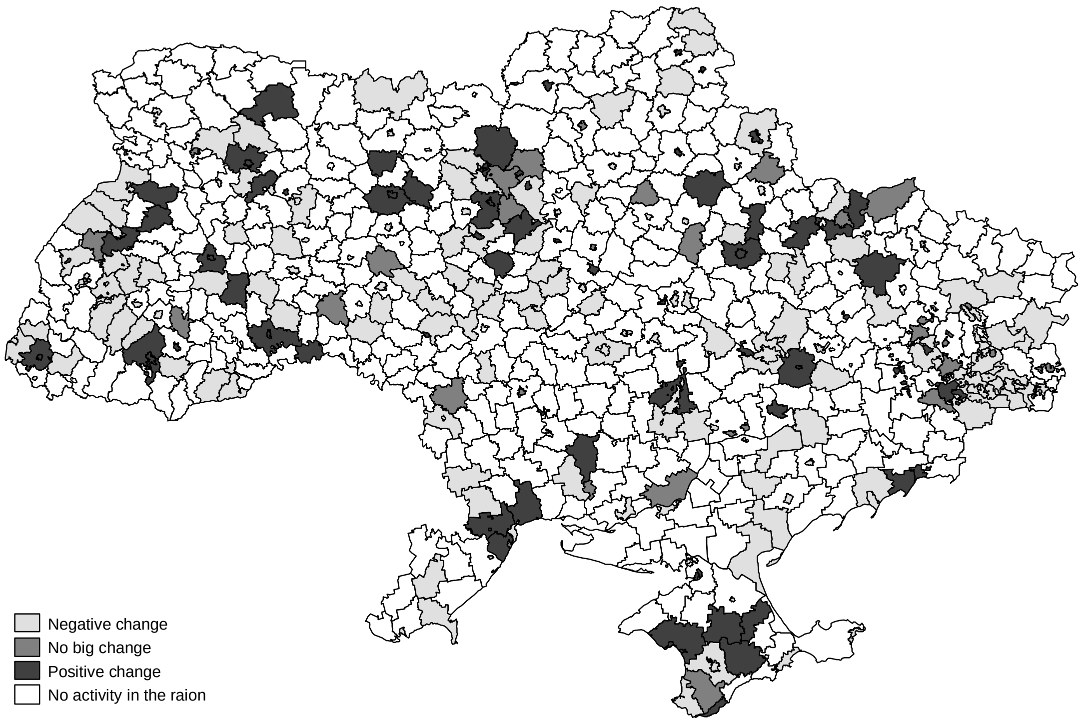


Figure 3.10, Spatial Distribution of Plants in Machinery, 2001 and 2005

Change in Machinery Employment, 2001-2005



Change in Machinery Number of Plants, 2001-2005

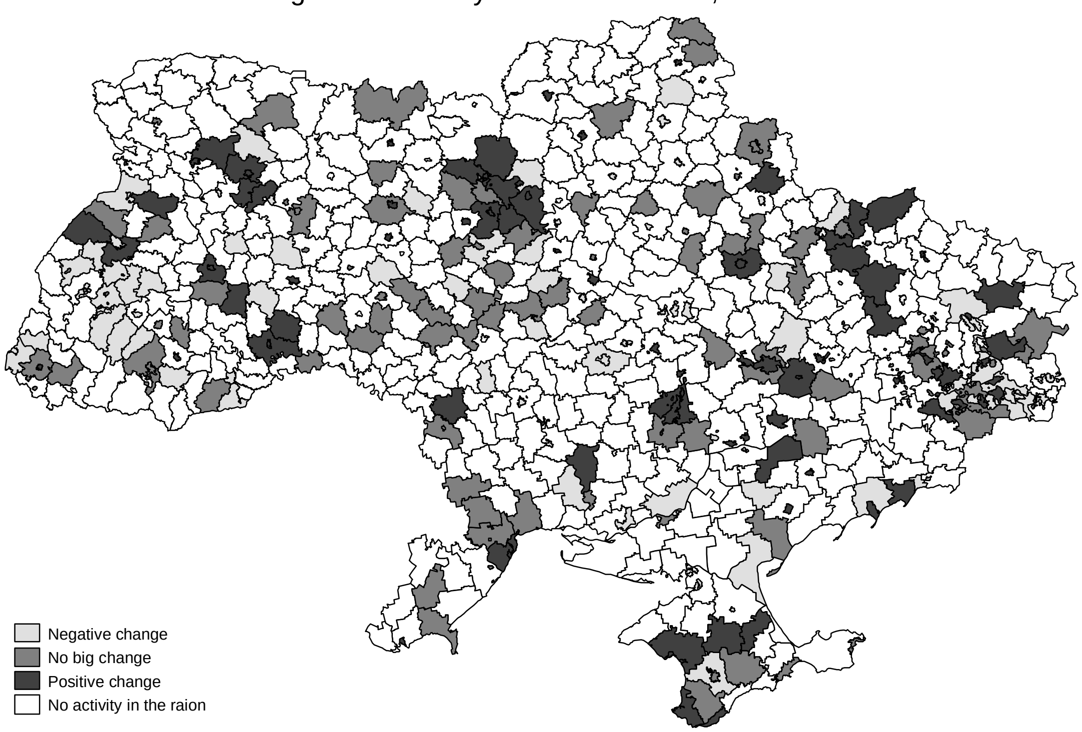


Figure 3.11, Changes in Machinery Employment and Number of Plants, 2001 – 2005

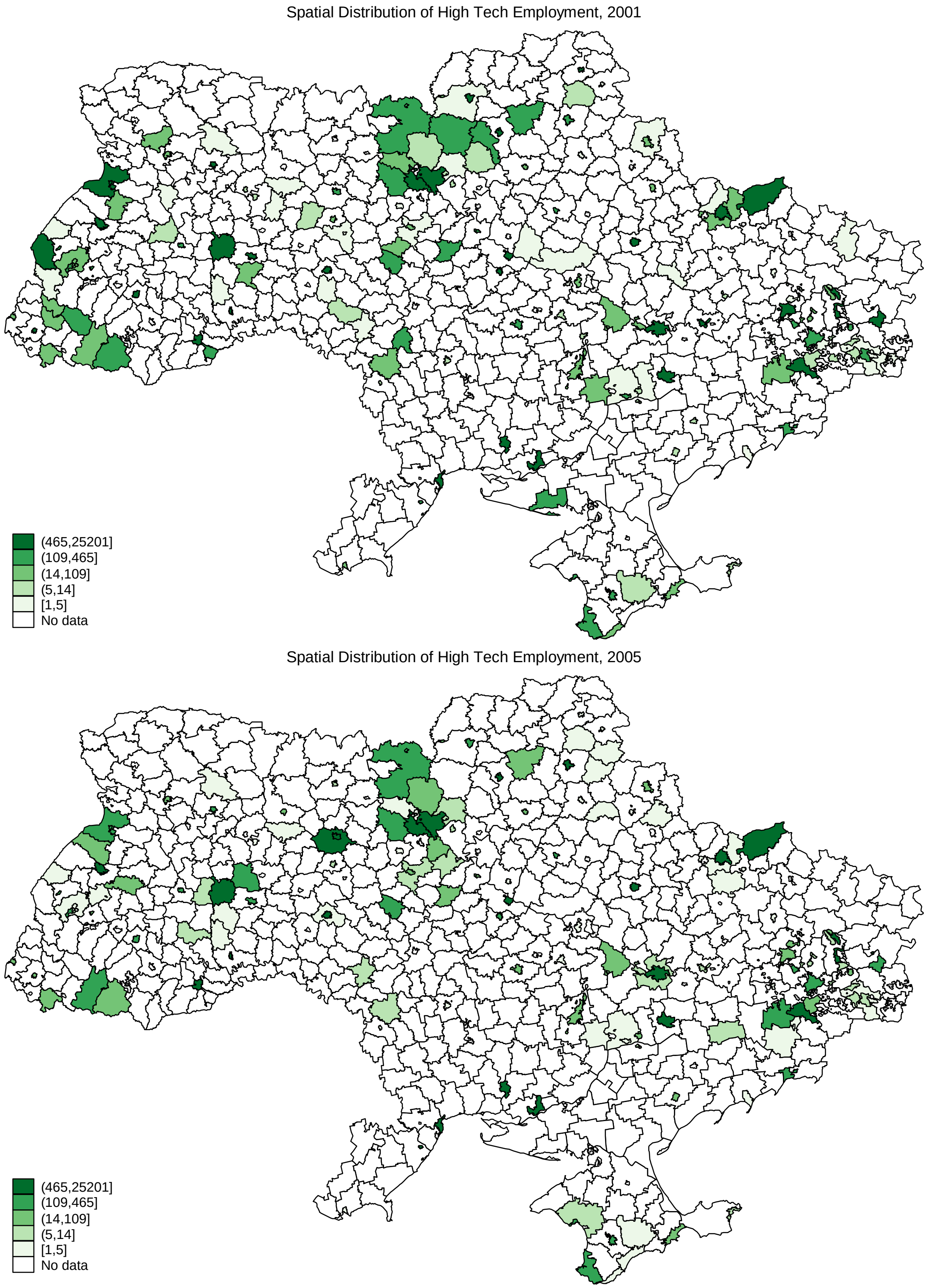
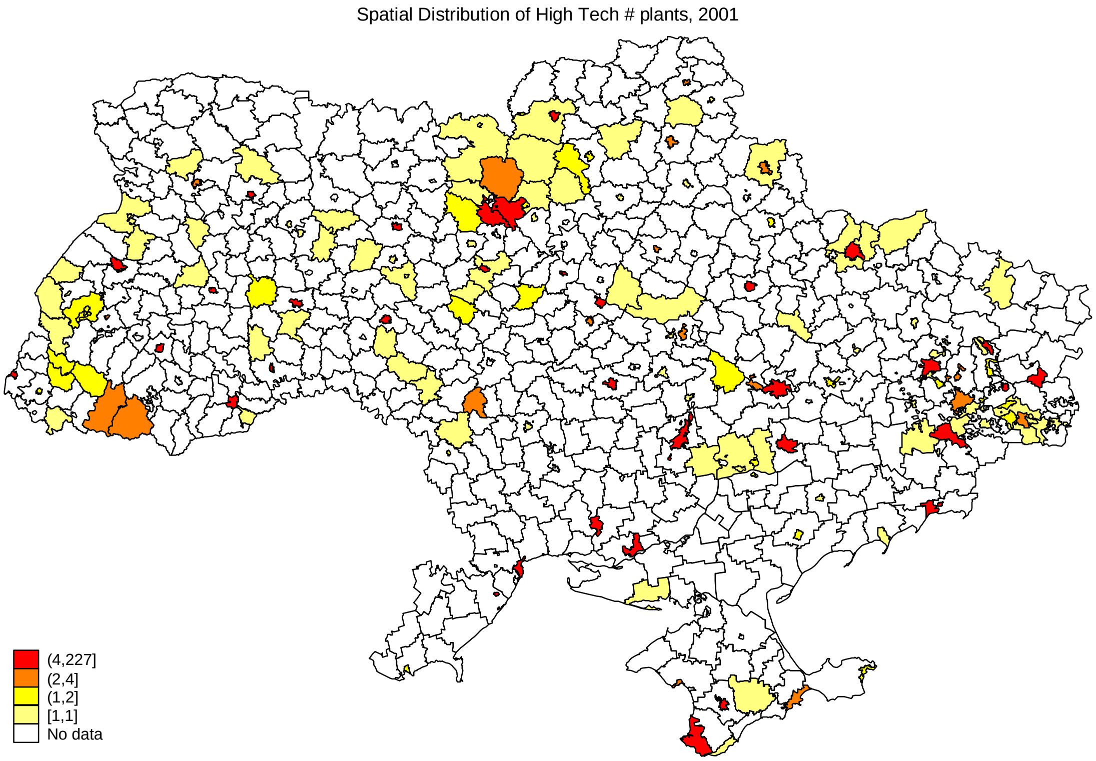


Figure 3.12, Spatial Distribution of Employment in High Tech, 2001 and 2005



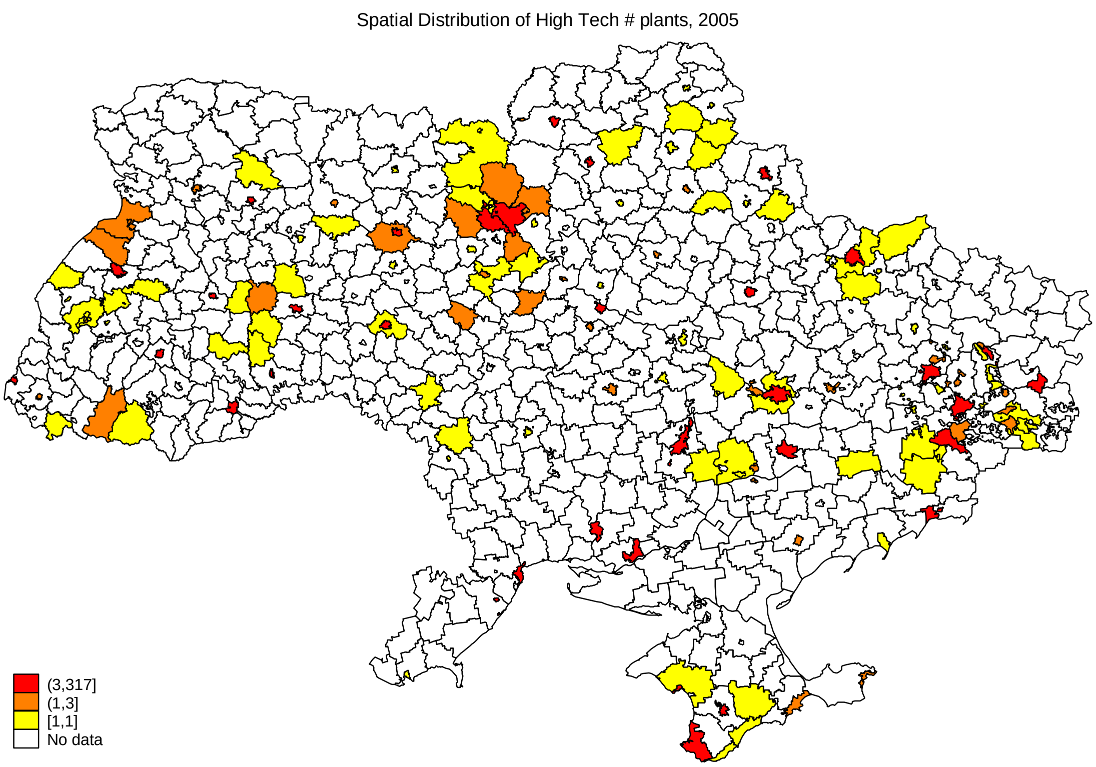


Figure 3.13, Spatial Distribution of Plants in High Tech, 2001 and 2005

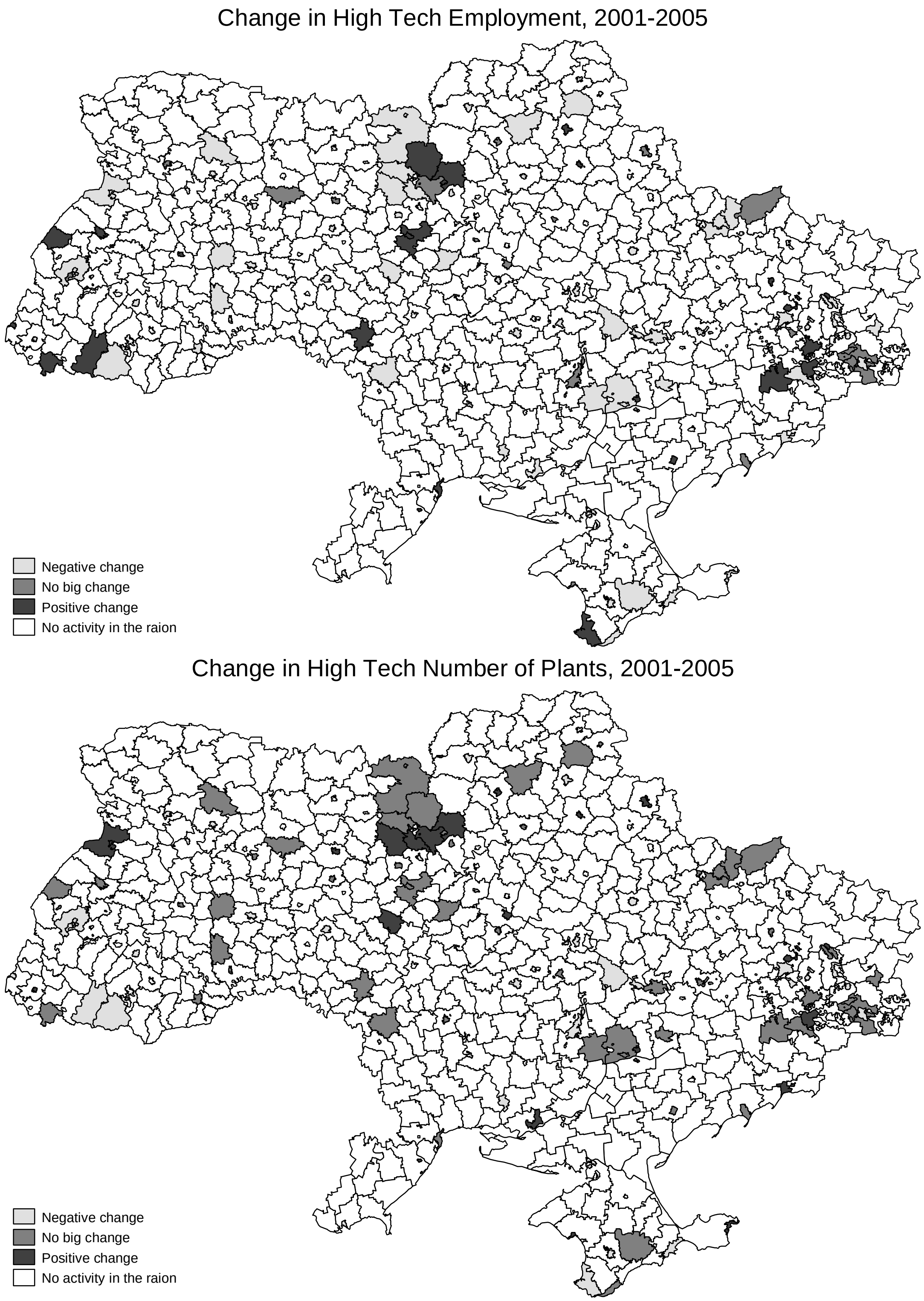


Figure 3.14, Changes in High Tech Employment and Number of Plants, 2001 – 2005

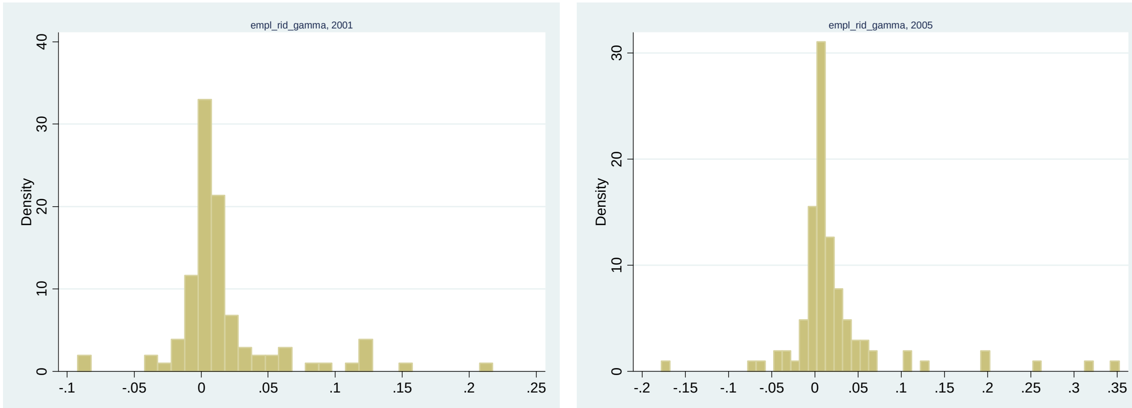


Figure 3.15, Histogram of Employment – Based γ, 2001 and 2005 (103 Three-Digit Industries).

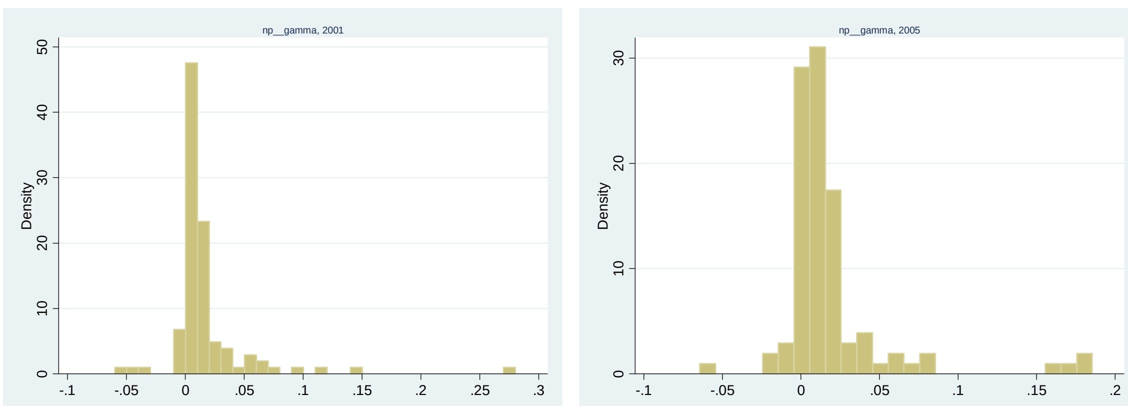


Figure 3.16, Histogram of Plant – Based γ, 2001 and 2005. (103 Three-Digit Industries

Chapter 4. Agglomeration Economies in Machinery and High Technology Industry Groups

4.1. Introduction

In this chapter I estimate agglomeration economies in two manufacturing groups, machinery and high-tech, in Ukrainian economy. The production function framework is a standard approach to estimate the agglomeration effects. In the most common setting (as mentioned, for example, in Henderson (2003)), agglomeration economies are assumed to be Hicks-neutral (not affecting a mix of production inputs). In the first major review on the topic, Eberts and McMillen (1999) propose the generic functional form:

$$y_j = g(A_j)f(x_j), \tag{4.1}$$

where y_j is firm's *j* output, $g(A_j)$ is the agglomeration component, and x_j is a set of production factors. In the later review, Rosenthal and Strange (2004) expand the agglomeration component arguing that all types of relationships between firms should be accounted for. They propose the following "benchmark" model of the total effects from agglomeration economies enjoyed by the firm *j*:

$$A_j = \sum_{k \in K} s(x_k, x_j)\theta(d_{ij}^G, d_{ij}^I, d_{ij}^T) \tag{4.2}$$

The first component of the sum $s(x_k, x_j)$ models all relationship between the firm *j* and another firm $k \in K$, where *K* is a set of all firms in the economy. Rosenthal and Strange (2004) also argue that the strength of the relationship and its possible effect on the firm's productivity attenuates with the distance. The second component of the model measures attenuation of agglomeration effects with distance, where distances between firsm *j* and *k* are measured in the geographic (d_{kj}^G), industrial (d_{kj}^I), and temporal (d_{kj}^T) space. The authors argue that intensity of relationships in either dimension directly affects spillovers between firms and their productivity.

Based on data availability, different authors took a number of approaches to estimate agglomeration economies in specification (4.1). Early studies estimated production functions for different industries separately based on data aggregated to the level of Metropolitan Statistical Areas (see Quigley (1998) for the review). For example, Sveikauskas (1975) used CES specification of the production function of manufacturing industries and found that doubling the city size increases industries productivity by six to seven percent, which suggests urbanization economies. Segal (1976) estimated the production function for all sectors rather than manufacturing only and found that this effect was about eight percent. Moomaw (1981) criticized both Sveikauskas for having capital omitted from the estimation and Segal for aggregating both manufacturing and service sectors. Moomaw (1981) found an alternative estimate of the city size effect which was less than three percent. Nakamura (1985) applied a translog specification, which is a second order approximation of the production function in Taylor expansion, and estimated not only urbanization, but also localization effects at the level of two digit industries for Japanese prefectures. Urbanization was proxied by total employment in the city, whereas localization was proxied by employment in the industry. He found that doubling industry size leads to productivity increase of about 4.5%, whereas city size effect is smaller, about 3.4%. His findings favor localization economies more. Similar results show up in Henderson (1986) comparative study of USA and Brazil, but Henderson did not find significant urbanization effects. On the other hand, Moomaw (1983) finds evidence of both localization and urbanization economies. To summarize, aggregation of the data at the level of two-digit industry groups and non-availability of key variables (such as land or internal production) failed to lead to robust conclusions in favor of urbanization or localization economies alone, but nonetheless they have shown that agglomeration economies do have effect on firms' productivity.

In recent years, firm level data has become increasingly available. Glaeser et al (1992) explored growth determinants for six largest industries within an MSA. They found that more heavily concentrated industries grow slower, which contradicts predictions of MAR externalities. Another important finding is that greater variety of non-dominant industries in the city helps own industry grows, which is consistent with Jacob's (1969) predictions. In a similar paper by Henderson, Kuncoro and Turner (1995), the result was less clear-cut: authors found that specialization (and subsequent concentration) may have positive effect for more matured manufacturing industries, whereas diversity is more preferable for the growth of younger high-tech industries. Both papers have shown that city characteristics may affect industry's grows for a very long period of time. Rosenthal and Strange (2003) considered effects of births of new establishments on employment at ZIP code level of geographical aggregation. They found that strong localization economies are present (new firms are more likely to be attracted to ZIP code areas where this industry is already represented), though they rapidly attenuate with the distance. Urbanization effects were found to be less robust and were likely to differ across various industries. Another important finding is the fact that "industrial structure and corporate organization" is important: "more entrepreneurial industrial system promotes growth."

Henderson (2003) also used firm-level census-type data to estimate production function (4.1) directly for machinery and high-tech sectors using various measures for agglomeration economies. Localization economies were proxied by employment and plant counts in the same industry and same geographical area. Estimates of urbanization economy were based on various "lack of diversity measures", such as total manufacturing employment in the same MSA or degree of MSA employment specialization outside the own industry. The lagged effects are also measured to account for the effect of "past experience." He found strong localization economies for high-tech firms and almost no localization effects for machinery firms. He did not find significant urbanization effects for any of the subgroups of industries. Another interesting finding in this paper is that proximity of firms seems to be more important than simply an effect of employment in the same industry. The author comes to a conclusion that possible source of localization economies is in existence of firms per se rather than in the related employment scale.

Rosenthal and Strange (2004) believe that the approach taken in Henderson (2003) is the most appropriate way for a productivity-based study of agglomeration. I follow this approach to analyze external effects for Ukrainian industries and compare the results with Henderson's. I also extend Henderson's finding by analyzing the effect of the ownership type on external economies within and outside of a firm.

4.2. Estimation strategy and issues

I follow the methodology proposed by Henderson (2003). A standard log-log production function for a given establishment *j* in the area *m* in time *t*, the output y_{jt} is

$$\ln(y_{jt}) = \boldsymbol{\alpha} \ln(\mathbf{X}_{jt}) + \beta \ln \mathbf{E}_{jt} + \gamma \ln \mathbf{I}_{jt} + \delta_t + \phi_{mj} + \eta_{jt} + \varepsilon_{mjt} \tag{4.3}$$

where $\mathbf{X}_{jt}$ is a vector of production inputs, E_{jt} is a vector of agglomeration variables, $\mathbf{I}_{jt}$ is a vector of institutional variables, δ_t is the time fixed effect; ϕ_{mj} is the location fixed effect; η_{jt} is the three-digit industry fixed effect, and ε_{mjt} is the error term.

The data in my disposal are annually submitted to the National Statistics Office by individual firms. When a firm has branches in a given year, it is the firm's responsibility to report data for all its branches. It is evident from the data that not all firms submit statistical forms annually, and some of them do not submit branches information as well. This conjecture is based on a number of gaps in the data within the same firm's ID. Such peculiarity of this dataset makes it difficult to ensure if the absence of a firm in the data may be attributed to a "true" exit or a simple lack of data. The same refers to new firms, since it is not obvious from the data when they started the operations. Brown and Earle (2006) also mention that some firms could have changed their identifiers during the study period. They performed a matching of "old" and "new" identifiers when they could make a conjecture that it is the same firm to reduce possible attrition and entry problem. However, the number of such matches was not significant (less than one percent of observations in a sample of large manufacturing firms between 1992 and 1998), whereas in some cases, it is very difficult to distinguish between a firm's identifier change and true exit of one firm and replacement of that firm with a new one, in the similar industry, at the same address and similar characteristics. Therefore measurement errors are possible when performing such a matching exercise.

Taking these considerations into account I decided to leave the sample as is, realizing that attritions and exits may affect the estimates of productivity. Olley and Pakes (1996) point out that implicit decisions of firms regarding a particular mix of inputs may lead to the simultaneity bias, whereas unobserved entry and exit factors may create a related selection problem. Both problems have been addresses in econometric literature, but no universal solution has been proposed. One of earlier solutions, balancing a sample, is not appropriate in this study for two reasons. First, the attrition rate is relatively high. For example, in machinery about a quarter of firms entering the sample in 2001 changed their major activity by 2005. In the high tech sector, the attrition rates a similar. Second, balancing the sample still leads to the upward bias in the productivity estimates. Exiting firms may have lower productivity in the periods prior to the exit, whereas entering firms may have higher productivity after several periods that the average in the sample. Olley and Pakes (1996) propose a novel algorithm based on the semiparametric approach to the estimation of the production function. They use investment as an instrument in their model. Levinson and Petrin (2003) argue that only a fraction of firms makes regular investments, and propose a similar procedure which utilizes the data on material costs. The data available for this study do not include reliable investment or material cost information, and hence this approach is not feasible. It is also important to compare my results with those in Henderson (2003) and thus follow his estimation strategy. Henderson chose a simple fixed-effects specification after experiments with 2SLS and GMM estimates, which he rejected because it was difficult to find appropriate instruments. I follow the same approach for the similar reasons.

I estimate the specification (4.3) separately for two industry groups: high-tech firms and machinery firms.

After each establishment was allocated to an industry-region cell I could calculate the agglomeration measures. Estimating *localization* effects, ideally I want to model all interactions between firms in the same industry and the same location. Marshall (1920) indicated that such interactions can take a form of knowledge spillovers, labor market sharing or input market sharing. As I can hardly obtain direct measures of all possible interactions, I proxy them through the labor market pooling (as measured by employment in the same industry in the same location, less own employment) and interfirm relationships (which is reflected in the establishment counts in the same industry/location cell, less own firm). I experimented with several "distances" in both "industrial" and "geographic" dimensions. As a preferred specification, I have chosen two: plant counts and employment in the same industry group (machinery or high-tech) and the same raion, and plant counts and employment in the same three-digit industry group (KVED-3) and the same QMSA. The industry group consists of several KVED-3 groups, and QMSA consists of several raions. That is why the first group of agglomeration variables measures relationships between firms in the wide industrial space, but emphasizes geographic proximity, whereas and the second group of measures accentuates "industrial" proximity loosing geographic distances between firms.

For *urbanization* effects, I tried to estimate diversity forces at work in a given area. Diversity reflects interactions between firms from different industries. The most standard measure of diversity is the total number of establishments and total manufacturing employment in the area, variables used by many authors in the precedent work (for example, in Svejkauskas (1979) and Nakamura (1985)). I used total employment (less own employment) and establishment counts (less the firm itself) in the raion or QMSA. Hederson (2003) also used these measures.

The set of institutional regressors consists of indicator variables "Urban" and "Subsidiary," as well as of variables indicating the ownership type. I compare two specifications: with and without ownership variables, to estimate the effect of agglomeration on owners' location decisions. When I estimate the ownership effect I use the cross-terms of agglomeration and ownership variables:

$$\ln(y_{jt}) = \boldsymbol{\alpha} \ln(\mathbf{X}_{jt}) + \beta \ln \mathbf{E}_{jt} + \boldsymbol{\Gamma}_1 \ln \mathbf{I}_{jt} + \boldsymbol{\Gamma}_2 \ln \mathbf{I}_{jt} \mathbf{E}_{jt} + \delta_t + \phi_{mj} + \eta_{jt} + \varepsilon_{mjt} \quad (4.4)$$

The cross-effects of agglomeration and ownership will be shown in $\boldsymbol{\Gamma}_2$ group of regression coefficients.

I use location fixed effects to control for location specific amenities which might simultaneously improve productivity and attract many firms. Ideally I would prefer to use plant specific fixed effects as in Henderson (2003) and Brown et al. (2006). However, a short time span in my data does not allow for substantial variation in the agglomeration variables. On this basis, I hope that location (particularly, QMSA) fixed effects with urban indicator variables will be able to absorb most differences between the firms. QMSA fixed effects account for specific transportation conditions, business climate, or particular local amenities. The "urban" variable in the estimation of agglomeration effects at the level of a KVED-3 group within a QMSA also accounts for specific urban characteristics. The agglomeration measures are the same for all firms within the industry-region cell, which also justifies the use of location fixed effects rather than firm fixed effects. Location specific amenities simultaneously affect productivity of all firms within the cell. Hence, the regression coefficient on the agglomeration variable indicates an average effect for all firms.

Given that every group is composed of several three-digit industries, I also experimented with industry fixed effects. The group of machinery firms is relatively more homogeneous (it is composed by three-digit industries from the same two-digit industry sector), and I do not expect that coefficients will differ significantly for industry fixed-effect specification. The high-tech group, on the contrary, is heterogeneous both in terms of industry composition (three-digit industries constituting the group belong to different two-digit industrial sectors), and, possibly, production or business processes. If there are any additional relationships between these three-digit industries, it is possible that coefficients in the industry fixed effect specification will reveal them by behaving differently. I estimated specifications both with and without industry fixed effects and found no significant difference in my coefficients of interest. In my further work I present only results with industry fixed effects.

4.3. Estimation results and discussion

4.3.1 General Comments

The estimation results of the model (4.3) for the machinery sample are presented in Table 4.1, and the high tech sample results are found in Table 4.2. Production factor elasticities are consistent with those found in other agglomeration studies, as well as with productivity studies of transition economies. The coefficients have expected signs and are strongly significant. Henderson (2003) reported capital elasticities in the range of 0.03 to 0.07 in machinery and 0.05 to 0.08 in the high-tech industry groups. Using Ukrainian data, Brown et. al. (2006, unpublished appendix) estimated capital elasticities of 0.094 in "Machinery and equipment" and 0.044 in the "Electrical and optical equipment" sector, which is similar to high-tech industry group in my study. I estimated capital elasticities in the range of 0.08 to 0.088 for machinery and 0.108 to 0.121 for high-tech. In general, the relationship between the industry groups resembles that Henderson's. My labor elasticities are about 0.94 in machinery and between 0.96 and 0.99 in the high-tech. Brown et al (2006) used a similar dataset and also lacked material cost data; they report labor elasticities slightly above one in both industry groups. Henderson's results are half as large compared to ours, but he was able to include the material costs. In general, it is likely that my labor coefficients may include the effects of other omitted factor variables.

The sum of my production factor elasticities (capital and labor) slightly exceeds one both for machinery and high-tech indicating close to constant returns to scale. I tested the hypothesis that $\alpha_K + \alpha_L = 1$ and failed to reject it in all equations. This result contradicts somewhat to Henderson's who reported the sum of factor coefficients to be slightly below one. The discrepancy may be attributed to the lack of the material data or measurement errors. On the other hand, the returns to scale in my work are similar to those found by Brown et al. (2006) with the Ukrainian data.

The coefficient on the subsidiary status is always negative and strongly significant. A possible cause for the negative effect is the fact that subsidiaries as parts of broader intrafirm networks are located in the places determined by their mother companies or headquarters and do not always reflect the reasoning of an independent profit-maximizing firm. I ran separate estimations for subsidiaries and non-affiliated ("independent") firms. For independent firms, there were no noticeable changes both in the factor elasticities and agglomeration coefficients. For affiliated firms, the capital elasticity fell sharply to below 0.01 or even negative and in most estimation equations turned statistically insignificant, whereas the labor elasticity increased to above one. The agglomeration coefficients in most specifications have turned insignificant and sometimes changed their sign to negative. The rise of the standard errors was especially pronounced in the high tech sample. One of possible reasons for this may be a relatively low number of observations in the subsidiary subsample (about eleven percent in both industry groups on average), which could inflate the standard errors.

In specifications when agglomeration variables were measured at the level of three-digit industries in the QMSA, I included the "urban" variable to mark observations located in cities or towns within the QMSAs. The coefficient on this variable was positive and strongly significant both in machinery and high-tech samples indicating higher productivities of firms in urban areas. This result generally follows a theoretical finding that productivity in cities tends to be higher. It is possible that this variable partially measures the urbanization economies effects. However, inclusion or omission of this variable does not affect the value and the significance of the agglomeration effects in all specifications.

4.3.2 Localization economies

Even though I compare my results to those of Henderson (2003), the major focus of my paper is on the effect of ownership changes on external scale economies. I start with the basic specification first.

I found that localization economies are present both in machinery and high-tech groups at the level of the industry group in the raion for all four agglomeration measures. This result contradicts Henderson (2003) that strong agglomeration effects are present only in the high-tech group, but scale economies in machinery were found non-existent. The magnitudes of the coefficients are different, though. Doubling the number of plants in the same industry group in the same raion increases firms' productivity by approximately nine percent in machinery and seventeen percent in high-tech (compared to 2.3% and 8.0%, respectively, in Henderson's study). Doubling employment rather than the plant counts increases productivity by 7.3% in machinery and 11.7% in the high tech group. At the level of a 3-digit industry in the QMSA, the machinery group also reveals external scale effects, but the significance of the results is lower. The employment effect falls to 4.0%, whereas for the plant counts, the effect is 7.4% and insignificant. In the high-tech group, coefficients are surprisingly negative (at –5.1%) for employment and also negative (–9%) and statistically insignificant for the plants counts. A possible reason for this sign change may pose questions about robustness of the localization coefficient at this level of agglomeration or the way the high tech sample is composed. On the other hand, aggregation of the data at the level of a three digit industry group in the QMSA assumes shorter distance between firms in the industrial space, but greater physical distances. The evidence of "de-agglomeration," or external scale "diseconomies," may indicate overcrowding of the area with the firms in a particular three digit industry and signal about the need for a further diversification. Yet another possibility is a large size of QMSA's and hence weaker linkages between firms, on average. In general, the issue of the market potential as well as the optimal number and size of firms in the region (and of the regions themselves) remains open to a further investigation.

Another important observation is relevant here. My results support Henderson's conclusion that it is the number of plants rather than employment in the same industry that provides localization effects. Intuitively, this result suggests that knowledge spillovers occur between entrepreneurs (firm owners) rather than between employees of the firms, and thus even though employment may play a role in creating a common labor pool, it is not the primary reason for firms to concentrate in a particular location, at least in these particular industry groups. I will return to this issue when I discuss the ownership effects.

I also experimented with the quadratic agglomeration terms assuming that localization economies may start to decrease after the number of firms or employment in an industry-region cell reaches a certain point. In all cases, the coefficient by the quadratic term turned out insignificant, suggesting a linear relationship between concentration of economic activity and firms' productivity.

In my previous specification I explicitly assumed the same slope of agglomeration coefficients in all years. However, transition processes have proceeded fairly quickly, and even annual changes in the business environment could be significant. I estimated separate models for each pair of consecutive years between 2001 and 2005. The results are present at Table 4.3 and graphically at Figure 4.1 for machinery and Figure 4.2 for the high tech sample. The tables show results of sixteen separate regressions grouped by years and the agglomeration measures.

Due to smaller sample sizes and possible lack of considerable variation in the agglomeration variables, results are insignificant for some periods. Nevertheless, the plant count measures are generally greater than the employment measures, again suggesting that spillovers are more likely to occur between firms rather than between employees of different firms.

Another observation is that the effects seem to be greater in the earlier years in the sample and descend towards 2005. For example in machinery, the effects of the plant counts in the same industry group in the raion fell from 9.6% in 2001-2002 period to 7.9% in the last period. In the high-tech, it the same effect increased from 20.4% in 2001-2002 to 21.7% in the next period and fell to 12.0% in the last period. One of possible explanations is that the market is becoming more satiated, and hence the marginal effect of new firms entering the market is lower. Another explanation is that with the development of the economy, my results converge to those of Henderson (2003)

It is also remarkable that coefficients for most periods for both types of agglomeration variables at the level of three-digit industry in QMSA are statistically insignificant and some have negative signs. This may suggest the lack of agglomeration economies between firms in the QMSA, most likely, due to greater distances between firms.

4.3.3 Ownership effects

One of the most important features of the Ukrainian economy is a relatively large number of firms in the state property or formerly owned by the state. Their geographical location was predetermined, and for the privatized firms, their industrial activity and, in some cases, even employment had to be preserved into several years after the privatization. In the contrast, new private firms have more discretion in the choice of place and the use of production factors. Even though they are unable to change their physical location quickly after setting up they are more likely to relocate to another sector of the economy where they expect higher profits. Redistribution of production assets between sectors may affect external economies for all firms in the region, both private and public.

My data confirm lower mobility of state firms. Out of 3,042 plants in machinery in 2001, there were 179 public and 2,863 privately owned. Of them, 1,994 were operating through all five years until 2005, 82 public and 1,912 private. Over this time span, 493 private firms, or almost a quarter, switched from machinery to another activity within manufacturing, of which 311 plants (sixteen percent) switched to a completely different sector of the economy. Among public firms, twenty plants (which also constitute about a quarter) switched from machinery to another manufacturing activity, but only seven plants (eight percent) switched to a completely different sector.

Out of 1,010 high-tech plants in 2001 (90 public and 920 private), 694 (51 public and 643 private) operated through 2005. Among the private firms, 237 plants, or thirty seven percent, switched from high-tech to a different manufacturing activity, and another twenty seven percent switched to a completely different sector. Among the public firms, ten plants, or one fifth, switched from high tech to another manufacturing industry, but only two plants (four percent) switched to a completely different sector of the economy. Hence, my data reveal much lower inter-sector mobility for publicly owned firms.

I have amended my dataset with the pair of the ownership variables (indicating primarily domestic and primarily foreign ownership, the majority state ownership being the base category) and corresponding interaction terms with the agglomeration variables, and estimated model (4.4) again. The results (presented in the same tables as the basic specification) indicate that ownership type matters for external scale economies: privately owned firms enjoy greater effects compared to the state owned firms, whereas foreign firms experience larger effects than primarily domestic firms.

Introduction of the ownership variables has slightly increased the factor elasticities. The coefficients by the ownership variables (DO and FO) are statistically significant and have the expected signs. Productivity of domestically owned private firms is distinctly greater compared to the state firms, whereas for the primarily foreign firms, the effect further increases approximately twofold in both industry groups. It is difficult to immediately compare the results with other studies. Brown et. al (2006) who used a very similar dataset, used a series of lagged and forwarded ownership values to estimate the effects of the change in the property type whereas I use the same period ownership type as simply another control. Nonetheless, the order and direction of the effect appear to be close to results at Brown et al (2006).

As a result of the introducing the cross terms, the total effect of agglomeration variables is now decomposed into the effect on the majority state firms, primarily domestic private and primarily foreign private firms.

We first explore the machinery sample. At the level of the industry group in the raion, the magnitude of the employment effect did not change for state owned plants, though the standard errors almost tripled. The effect for primarily domestic owned firms is nearly zero and insignificant, whereas primarily foreign firms enjoy 8.7% increase in productivity (though not statistically significant) totaling at 16%. For the plants measure, both state and primarily domestic coefficients are insignificant. In terms of magnitudes, the external scale economies fell to 5.9% for the state owned firms and, with an increase of 2.5%, totaled at 8.4% domestic firms. The effect of the primarily foreign ownership is additional 16.4% totaling at 22.3%.

At the level of three digit industries in the QMSA, the decomposed effect of employment in the same sector in the area has slightly decreased for the state owned firms from 4%% to 3.8%, but it shows a total increase to 4.6% for primarily domestically owned firms and increase to 13.4% for primarily foreign owned firms. At the plant level, the decomposition suggests that the agglomeration economies fall from 7.4% to 5.9% for the state owned plants, increases by 2.8% to 8.4% for primarily domestic plants and enjoys a considerable 16.1% increase totaling at twenty two percent for primarily foreign plants.

In the high tech sector, the results are similar. At the level of the industry group in the raion, the local employment effect decrease in magnitude to 4.3% and turns insignificant for the state owned firms, increases to 12.4% for primarily domestic firms and to 14.5% for primarily foreign firms. The coefficient on the number of plants also becomes insignificant and drops in value from 16.7% in the non-decomposed sample to 6.9% in the decomposed sample for the state owned firms. For primarily domestic firms, an increase is 10.7% to 17.6% total, and primarily foreign firms also enjoy an increase by 16.2% to 23.1% total. At the level of three-digit industry sector in QMSA, the local employment effect remains negative and significant for the state owned firms, but the absolute magnitude changes from 5.1% to 8.4%. The marginal increase of the agglomeration effect for domestic and foreign firms is 3.7% and 3.9% respectively. Since only state firms indicate a negative effect of local industry employment increase, we may observe a possible case of the "de-agglomeration effect" for the state firms. In other words, state owned firms may be loosing their competitive edge at the presence of the private firms in the high tech industries. The effect of the local plants is even more pronounced. The magnitudes plummet from negative 9% down to negative (and significant) 23% for the state firms whereas private firms show marginal increases by 15.3% for domestic and 20.7% for foreign firms. Even though the total effects still remain negative, it is most probably state firms that drive the total results down.

In general, we observe a lower agglomeration effect for the state owned firms for all measures in both industry groups, with different levels of statistical significance. The agglomeration effect for domestically owned firms is not substantial in machinery and positive and significant in high tech. The primarily foreign firms seem to enjoy the largest effects of agglomeration in both industry groups. Even though "cherry-picking" for the foreign firms is possible (in a sense that foreign owners initially choose the most productive assets when consider opening their business in Ukraine), I still may conclude that they manage to choose the most productive locations and enjoy the agglomeration effects to the largest extent. My estimations also suggest that ownership or firm management is a more important channel of agglomeration than employment, since it is the number of firms in vicinity that brings about higher magnitudes of plants productivity increase. Another observation is the difference in statistical significance of the results at the three digit industries in the QMSAs between machinery and high tech. It is possible that geographic distances between firms in machinery play a greater role in accruing the agglomeration effects than distances in the industrial space, and therefore effects at this level of aggregation are insignificant. In the high tech industry group, on the contrary, we observe lower standard errors for agglomeration coefficients and greater significance. One of possible explanations is the difference in the number of firms between the groups. Since the high tech sample is only a third of the machinery, increasing the number of high tech firms in the neighborhood, even widely defined (such as QMSA), still affects firms' productivity. Greater heterogeneity of the high tech sample may also add to the effect. The number of firms in the same three digit industry in the area is relatively small, and an additional firm plays a greater role for productivity than in machinery.

4.3.4 Urbanization economies

Rosenthal and Strange (2004) refer to extensive literature on the effect of urban diversity on the firm productivity. Traditional understanding of this source of external scale economies was developed in Jacobs (1969) and assumes information spillovers between firms from various industries, lower search costs, "cross-fertilization" among industries, better quality and access and public infrastructure, and possibility to outsource standard routine activities to other specialized firms. Henderson (2003) mentions that productivity studies in manufacturing provide only weak empirical support of urbanization effects. One of possible reasons is the lack of mutual agreement on how to measure urban diversity. Henderson (2003) proposed a whole range of "metropolitan specialization" measures, or "non-diversity" indices, but did not find any significant effect on productivity. He explained this by insufficient variation in these indices over time. However even after he has turned to general scale measures such as the number of plants and overall employment in the area he still did not find significant urbanization effects.

I will follow Henderson's approach and also estimate the urbanization economies using simple scale measures. I have calculated the total number of firms and employment in the raion and in QMSA (less own firm). Findings are presented in Table 4.4. For both industry groups, the QMSA measures (columns 1, 2, 5, and 6) turned out to be insignificant and have a negative sign. This result is comparable to Henderson's findings who also estimated the effects at the MSA level. The raion level measures, on the contrary, have shown significance both for employment and the number of plants. The difference between number of plants and employment is rather small within an industry group. Thus, doubling total employment in the raion increases a productivity of a machinery firm by 11.2%, whereas doubling the number of firms in the raion increases productivity of machinery firms by 10.1%. In the high tech industry group, the productivity increases by greater magnitudes: the effect of employment increase is 15.3%, and the effect of the plant count increase is 14.9%.

I believe that QMSA measures are insignificant in the first place due to local nature of urbanization effects: firms seem to have more business and customers interactions in the same city where they are located.

Even though my data show positive urbanization economies, I cannot know for sure the exact source of these effects. Henderson (2003) also found urbanization effects, but only in the machinery group. In his high tech group, his urbanization effects turned out to be statistically insignificant. In my sample, significance is present in both sectors. Moreover, in high tech, the effects are greater in absolute value. It may indicate that high tech firms are more reliant on other businesses in the area. My data indicate that about ninety three percent of high tech firms are located in cities compared to eighty eight percent in machinery. This may suggest that high tech firms are more reliant on particularly urban infrastructure and amenities. On the other hand, it may also indicate under-development of subcontractor relationships in machinery. As a result, most operations that are usually outsourced in the West are most likely to be performed "in the house" in Ukraine. The issue may be tackled by looking at co-agglomeration measures for machinery firms and presently is beyond the scope of this paper. In general, the effect of urban diversification on a firm productivity remains open and is subject to further analysis.

4.4. Conclusion

In this chapter I analyzed localization and urbanization economies in two Ukrainian manufacturing industry groups, machinery and high tech. Localization externalities arise from interactions between firms in the same industry in the same area. I used two levels of aggregation of the industry data: at the level of an industry group in raions and at the level of three-digit industries in "quasi-metropolitan areas." I found strong localization effects in machinery and mixed evidence in high tech. Both sectors indicated stronger effects at the level of industry group in the raion, which is comparable to Henderson's (2003) results. An important finding is the fact that it is plants rather than employment that drive the localization effects, which presumes that interactions between firms are channeled through owners and management rather than employees. This argument was sustained also in the dynamic decomposition of the results into pairs of years between 2001 and 2005.

This work contributes to the literature by exploring the effect of ownership type on the agglomeration economies. Since Ukraine as a transition country is characterized by a relatively large share of formerly and presently state owned firms, agglomeration economies may have different effects for firms with various ownership types. I decomposed the agglomeration effect into three streams according to the ownership type of the firm. Indeed, the state owned firms, being the least flexible in relocating their production factors to different uses, benefit the least from the agglomeration economies. The primarily domestic private firms in machinery do not indicate substantial use of agglomeration economies, but still manage to perform better than public firms and benefit more from agglomeration effects. In the high tech group, domestic firms are able to benefit more from agglomeration. The maximum effect is achieved by primarily foreign firms in both industry groups. Even though foreign owners could have chosen firms based on their performance, they still indicate higher capabilities in using the external economies than both majority state firms and primarily domestic private firms. Again, the effect is most pronounced at the level of an industry group in the raion, and number of plants plays a greater role in transferring the localization economies than employment.

Urbanization economies arise from interactions of firms from various industries in the same area. I used two different measures of urbanization scale measures: at the level of raion and at the level of QMSA. Using a measure similar to Henderson's (total employment and the number of plants in the MSA) I did not find any urbanization effects in both industry groups. However, in the contrast with Henderson, I have found urbanization effects in both industry groups at the level of raion. This indicates that urbanization economies are highly localized.

The study raises several issues for the further analysis. First, the use of firm fixed effects may single out the effect of external economies on a particular firm. On the one hand, this will help remove perfectly valid argument that changes in the environment affect different firms within a group in different ways. On the other hand, the span of time in my data is too short to allow for significant changes in agglomeration variables for each firm, and this may inflate the standard errors and ultimately render the coefficients insignificant. The best solution is to obtain a longer data series. Another issue deals with more appropriate measure of urbanization economies. According to Jacobs (1969), urbanization economies are channeled through diversity of firms and innovation processes within a city. The traditionally used measures of urban economies (amount of a certain activity in a city) do not capture this diversity to a full extent. For example, a city with multiple firms in the same industry and a city with the same number of firms in many different industries will present the same scope of urbanization economies if one uses the traditional measures, whereas the latter city is far more diverse and may provide more services to all firms located within its borders. Investigation into other possible measures of urbanization economies is another promising direction of the future research.

Tables

Table 4.1, Localization Economies and Ownership Effects in Machinery

	Localization effects				Localization and ownership effects			
	Group-Raion		KV3 - QMSA		Group-Raion		KV3 - QMSA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital)	0.071[a]	0.071[a]	0.066[a]	0.066[a]	0.076[a]	0.075[a]	0.071[a]	0.070[a]
	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)
ln(Labor)	0.939[a]	0.938[a]	0.945[a]	0.944[a]	0.945[a]	0.945[a]	0.951[a]	0.950[a]
	(0.026)	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)	(0.024)	(0.024)
Primarily domestic (DO)					0.676[a]	0.585[a]	0.695[a]	0.600[a]
					(0.089)	(0.111)	(0.085)	(0.205)
Primarily foreign (FO)					1.261[a]	0.677[b]	1.331[a]	0.800[c]
					(0.170)	(0.337)	(0.181)	(0.461)
Local empl. effect	0.073[a]		0.040[b]		0.073[c]		0.038	
	(0.017)		(0.017)		(0.043)		(0.052)	
Empl. + DO cross-effect					-0.006		0.008	
					(0.035)		(0.044)	
Empl. + FO cross-effect					0.087		0.095	
					(0.074)		(0.085)	
Local plants effect		0.092[a]		0.074		0.059		0.059
		(0.024)		(0.044)		(0.054)		(0.091)
Plants + DO cross-effect						0.025		0.028
						(0.041)		(0.065)
Plants + FO cross-effect						0.164[c]		0.161
						(0.092)		(0.124)
Subsidiary	-0.483[a]	-0.476[a]	-0.482[a]	-0.481[a]	-0.349[a]	-0.345[a]	-0.342[a]	-0.339[a]
	(0.057)	(0.055)	(0.063)	(0.063)	(0.051)	(0.050)	(0.054)	(0.055)
Urban			0.290[a]	0.290[a]			0.289[a]	0.289[a]
			(0.074)	(0.074)			(0.074)	(0.074)
Industry-Time f.e.	yes	yes	yes	yes	yes	yes	yes	yes
Observations	13028	13028	13352	13352	13028	13028	13352	13352
Number of msa60	56	56	56	56	56	56	56	56
R-squared	0.63	0.63	0.63	0.63	0.64	0.64	0.64	0.64

Standard errors in parentheses
[a] p<0.01, [b] p<0.05, [c] p<0.1

Table 4.2, Localization Economies and Ownership Effects in High Tech

	Localization effects				Localization and ownership effects			
	Group-Raion		KV3 - QMSA		Group-Raion		KV3 - QMSA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital)	0.117[a]	0.115[a]	0.108[a]	0.108[a]	0.123[a]	0.121[a]	0.114[a]	0.113[a]
	(0.035)	(0.036)	(0.038)	(0.039)	(0.034)	(0.035)	(0.039)	(0.038)
ln(Labor)	0.962[a]	0.960[a]	0.963[a]	0.962[a]	0.996[a]	0.995[a]	0.992[a]	0.992[a]
	(0.036)	(0.036)	(0.043)	(0.042)	(0.028)	(0.027)	(0.036)	(0.035)
Primarily domestic (DO)					0.535[a]	0.213	0.585[a]	0.170
					(0.192)	(0.237)	(0.211)	(0.242)
Primarily foreign (FO)					1.024[a]	0.500	1.106[a]	0.512
					(0.334)	(0.718)	(0.208)	(0.603)
Local empl. effect	0.117[a]		-0.051[b]		0.043		-0.084[b]	
	(0.015)		(0.022)		(0.033)		(0.040)	
Empl. + DO cross-effect					0.081[b]		0.037	
					(0.035)		(0.037)	
Empl. + FO cross-effect					0.102		0.039	
					(0.134)		(0.145)	
Local plants effect		0.167[a]		-0.090		0.069		-0.230[b]
		(0.032)		(0.065)		(0.049)		(-0.094)
Plants + DO cross-effect						0.107[a]		0.153[a]
						(0.039)		(0.051)
Plants + FO cross-effect						0.162		0.207
						(0.134)		(0.142)
Subsidiary	-0.490[a]	-0.490[a]	-0.471[a]	-0.471[a]	-0.359[b]	-0.363[b]	-0.350[b]	-0.348[b]
	(0.112)	(0.111)	(0.118)	(0.117)	(0.142)	(0.139)	(0.156)	(0.151)
Urban			0.543[a]	0.546[a]			0.550[a]	0.553[a]
			(0.112)	(0.111)			(0.106)	(0.106)
Industry-Time fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	3949	3949	4036	4036	3949	3949	4036	4036
Number of msa60	48	48	48	48	48	48	48	48
R-squared	0.62	0.62	0.60	0.60	0.63	0.63	0.61	0.61

Standard errors in parentheses
[a] p<0.01, [b] p<0.05, [c] p<0.1

Table 4.3, Year to Year Changes in Localization Coefficients

Machinery Group	2001-2002	2002-2003	2003-2004	2004-2005
ln (Employment in GRP/RN)	0.0785***	0.0869***	0.0766***	0.0626***
	(0.021)	(0.018)	(0.018)	(0.018)
ln (# plants in GRP/RN)	0.0960***	0.113***	0.102***	0.0793***
	(0.030)	(0.027)	(0.025)	(0.023)
ln (Employment in KV3/QMSA)	0.0383	0.0538***	0.0495*	0.0231
	(0.024)	(0.019)	(0.027)	(0.023)
ln (# plants in KV3/QMSA)	0.0841	0.127**	0.101	0.0333
	(0.056)	(0.057)	(0.069)	(0.048)
Industry-Time Fixed Effects	yes	yes	yes	yes
High Tech Group	2001-2002	2002-2003	2003-2004	2004-2005
ln (Employment in GRP/RN)	0.136***	0.142***	0.147***	0.0839***
	(0.020)	(0.017)	(0.016)	(0.021)
ln (# plants in GRP/RN)	0.204***	0.217***	0.210***	0.120***
	(0.023)	(0.045)	(0.049)	(0.034)
ln (Employment in KV3/QMSA)	-0.0522*	-0.0349	-0.0381	-0.0497
	(0.027)	(0.031)	(0.035)	(0.036)
ln (# plants in KV3/QMSA)	-0.178**	-0.0198	0.0377	-0.0984
	(0.076)	(0.074)	(0.10)	(0.10)
Industry-Time Fixed Effects	yes	yes	yes	yes

Table 4.4, Urbanization Effects

	Machinery				High Tech			
	QMSA		Raion		QMSA		Raion	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Capital)	0.065[a]	0.065[a]	0.066[a]	0.066[a]	0.108[a]	0.107[a]	0.109[a]	0.109[a]
	(0.017)	(0.017)	(0.017)	(0.017)	(0.038)	(0.038)	(0.037)	(0.037)
ln(Labor)	0.944[a]	0.944[a]	0.947[a]	0.948[a]	0.957[a]	0.957[a]	0.954[a]	0.954[a]
	(0.025)	(0.025)	(0.025)	(0.026)	(0.041)	(0.041)	(0.041)	(0.041)
ln (Employment / QMSA)	-0.120				-0.535[c]			
	(0.13)				(0.28)			
ln (# Plants / QMSA)		-0.181				-0.247		
		(0.15)				(0.33)		
ln (Employment / Raion)			0.112[a]				0.153[a]	
			(0.033)				(0.035)	
ln (# Plants / Raion)				0.101[a]				0.149[a]
				(0.026)				(0.031)
Subsidiary	-0.476[a]	-0.476[a]	-0.465[a]	-0.461[a]	-0.459[a]	-0.458[a]	-0.457[a]	-0.460[a]
	(0.062)	(0.062)	(0.060)	(0.059)	(0.11)	(0.11)	(0.12)	(0.12)
Urban	0.295[a]	0.295[a]			0.518[a]	0.519[a]		
	(0.074)	(0.074)			(0.10)	(0.10)		
Observations	13456	13456	13456	13456	4180	4180	4180	4180
Number of msa60	56	56	56	56	52	52	52	52
R-squared	0.63	0.63	0.64	0.64	0.62	0.62	0.62	0.62

Standard errors in parentheses
[a] p<0.01, [b] p<0.05, [c] p<0.1

Figures

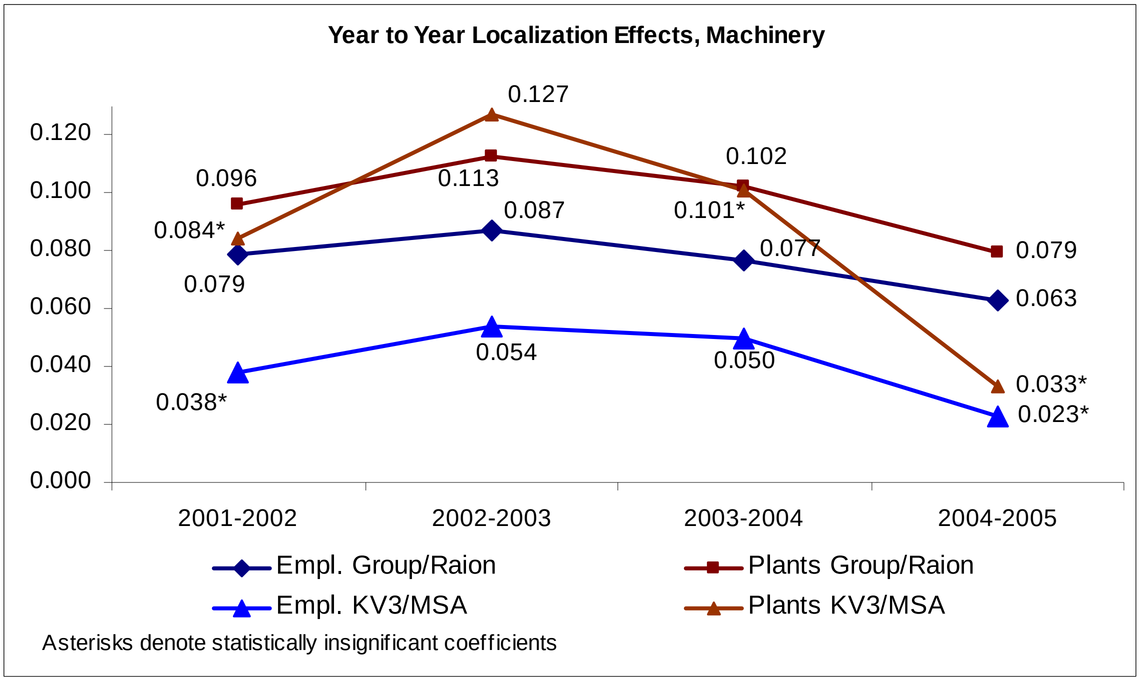


Figure 4.1, Year to Year Changes in Localization Economies in Machinery

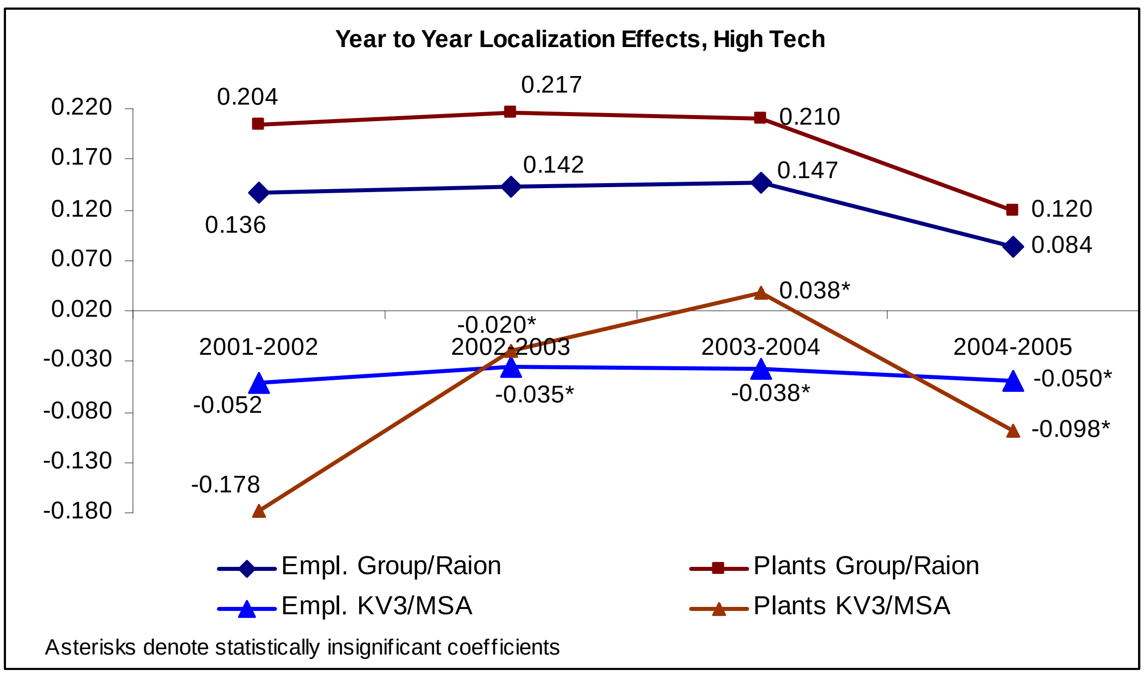


Figure 4.2, Year to Year Changes in Localization Economies in Machinery

Chapter 5. Attenuation of Agglomeration Economies

5.1. Measuring Agglomeration

Several studies show that productivity of firms increases in the presence of other firms in the same area and/or in the same industry. This effect is known as external scale economies, or agglomeration economies. There are two types of external effects. Localization economies are external to the firm but internal to the industry. Marshal (1890) proposed three major channels of such externalities: labor market pooling, input sharing and knowledge spillovers between firms. Urbanization economies are external both to the firm and the industry, but internal to the location of the firm. Jacobs (1969) introduced this type of externality pointing out a positive effect of the urban diversity. While Jacobs introduced urbanization economies as intuitively attributed to cities or at least to urban areas, there is no consent in the literature about the proper geographical and industrial scope of localization effects. In this chapter, I will concentrate on measuring localization economies at various levels of geographical and industrial aggregation. Using the same dataset of Ukrainian manufacturing establishments as in the previous chapters, I will demonstrate that the agglomeration effects attenuate with the distance, and a particular choice of the aggregation level may affect the estimated magnitude of the effects.

Quigley (1998) mentioned that estimation of the total factor productivity where agglomeration is introduced as a Hick-neutral parallel shift in the production function is the most standard methodology in the research of external scale effects. Eberts and McMillen (1999) review studies where this approach was taken. A particular functional form of the production function varies among studies, but most often Cobb-Douglas or generalized translog specifications are estimated. In earlier studies (for instance, Nakamura (1985)), constant returns to scale were assumed, whereas adherents of New Economic Geography (see Ottaviano and Thisse (2003) for a review) advocate modeling within CES framework (as proposed by Dixit and Stiglitz (1977)) with increasing returns. A unit of observation is usually an industry, a firm or an establishment.

There are a number of approaches towards measuring the agglomeration shift factor. Rosenthal and Strange (2004) proposed a “benchmark” model which defined an agglomeration component for each firm based on the relationships with other firms. They specifically mention that agglomeration economies tend to attenuate with distances between firms in geographic, industrial, or temporal space. The longer is a history of relationships between firms, the closer they are to each other geographically or in the production process, the grater is a probability that spillovers between them will take place and be more profound. The authors acknowledge that measuring exact distances is a hardly feasible task given the data requirements.

Measuring geographic distance is the most straightforward. When firms can be geo-coded, it is possible to construct firm specific distance measures. For example, Rosental and Strange (2003) used American firm-level data, whereas Cainelli and Lupi (2008) used Italian data. In both papers, a similar idea was used: authors built a series of concentric circles around each firm and measured agglomeration as the number of firms or employment in the own industry within circles of different sizes. Duranton and Overmann (2005) employed a different approach and developed a density function based on distances between all pairs of firms. All these studies have shown that agglomeration effects are strongest in the nearest vicinity and tend to attenuate after twenty to thirty kilometers.

When exact geo-coding information is not available other researchers have to partition geographic space into a discrete number of regions and assign each firm to one of the regions. The first approach assumes using small economic regions based on employees commuting patterns. Such areas were constructed for Italy (as used in Mion and Lafourcade (2003)) and France (Combes et al. (2008)). Another possibility is to use predefined regions as determined in the national territorial classifications. Examples of such regions range from villages and city districts to counties or states, provinces, prefectures, etc. When such regions are sufficiently small, they sometimes may be merged into greater conglomerates, like Metropolitan Statistical Areas in the US. The regional borders are politically or historically defined and do not necessarily account for particular location of economic activity. Head and Meyers (2004) provide an example of a watch-producing region which is divided not only by internal, but also national borders between France and Switzerland. Given the recent findings that localization economies tend to quickly attenuate, it is plausible to use a level of geographic aggregation that maximum internal distance within a region is close to twenty to thirty kilometers threshold. This corresponds to the level of NUTS-3 regions in most European countries and counties in the US.

Measuring the industrial distance between firms is more challenging. Firms may be assigned to an industry based on the similarity of their outputs, production processes, capital-labor ratios, or vertical relationships. However, the list of industries is strictly predetermined and nationally standardized. The hierarchy of levels of aggregation is imposed by the national industrial classification. Industrial classifications are similar across different countries, but not exactly the same. If a firm specializes in production of several outputs, the assignment process is based on the shares of those output types in the total sales. When researchers assign a firm to an industry, they do not have so much discretion as in the case of the physical space. The "industrial distance" between firms is assumed to be small if firms belong to different industries at lower level of aggregation, which together form same industry at higher level of aggregation. For example, firms from different three-digit industries but from the same two-digit industry are considered closer to each other in the industrial space than firms from different two-digit industries. Therefore localization effects are to a certain extent driven by the hierarchy of industries.

The issues of the minimal optimal industry size and industry boundaries are also very important. The major issue is to define a minimal industry size at which effects external to the firm but *internal* to the industry start to reveal themselves, but do not yet interfere with effects *external* to the industry as well. In the context of agglomeration economies this assignment makes sense only for sufficiently low levels of industrial aggregation. In other words, measuring economic distances correctly may override the national classification system and reshuffle small industry groups into larger ones based on a different principle. If it turns out that firms from two different small industry groups are similar in their production process, they should be assigned to the same "larger industry" group to which spillover effects are still internal, even though these firms may belong to different economic sectors in the standard classification. For example, an economic distance between concrete manufacturing and housing construction is supposedly very low, even though these two industries belong to different official sectors.

There were several attempts in the literature to measure economic distances between industries. One of possible ways to estimate it is to explore factor demands and sales between sectors using input-output matrices. Conley and Dupor (2003) constructed their own index measures of industrial distance for the US economy and found that sectors that share suppliers have higher correlation of growth rates. Ellison et al. (2007) used a simpler measure and found that input-output relationships between firms are important for co-agglomeration, or mutual location of firms from similar industries. A natural constraint for this approach is availability of the input-output matrices at sufficiently disaggregated level. Conley and Dupor (2003) used the matrices aggregated at two digits, whereas Ellison et al. (2007) got an access to the three-digit aggregation.

Nevertheless measuring both industrial and geographic distances is beyond the scope of most studies. Only a few datasets allow geocoding and exact measuring of physical distances between firms, as well as measuring industrial distances. Detailed input-output matrices are also quite rare. In practice, most datasets come from national statistics offices, when disclosure issues are applied. Therefore most researchers are bound with pre-determined industry-region cells and have a limited choice between a few levels of data aggregation.

Nevertheless even the analysis of aggregated data may lead to instructive insights about agglomeration effects within an industry and even about the attenuation speed. The agglomeration measure is the same for all firms within the cell and is typically measured as the total number of plants, overall employment or other size-related variable. The industry data are usually available at two or three digits, but sometimes four or even five digits are available for the majority of the firms. One-digit or even two-digit industries (sectors) are usually too large and heterogeneous, whereas four-digit industries look rather scattered on the map. For example, Combes et al. (2008) call three-digit aggregation as a "reasonable compromise." I will follow the same considerations. The level of territorial aggregation is usually also bound by two or three levels of administrative division of a country. The regions at the third level correspond to the European NUTS3 regions and comparable in size to the U.S. states. In Ukraine, the level of raions corresponds to this level of aggregation.

5.2. Residual agglomeration

In this chapter I define the minimal industry-area cell by aggregating the regional component at the level of a raion, and the industrial component at the level of a three-digit industry code. I will call this cell "primary." Given the data availability, this is the minimum level at which the localization economies take place. I compare the localization effects estimated for primary cells with more aggregated measures. I compare three series of estimates to baseline estimates using the primary measures. First, I expand the geographic area by measuring the agglomeration effect within the same three-digit industry at the QMSA level. Second, I will use a different geographical dimension and explore agglomeration effects within the same three-digit industry in the neighboring raions, which in a sense form a "spatial lag" of agglomeration economies. Then I expand the industrial dimension by measuring the agglomeration effect within the raion (smallest geography) at a higher level of industrial definition (the two-letter manufacturing subsector). Estimation is performed separately for each of eleven subsectors of the Ukrainian economy. For each expanded cell, I will estimate two models: one with a single agglomeration variable measured at the level of the expanded cell, and another with two agglomeration variables: the first measured at the level of the primary cell, and the second is the "residual" agglomeration, the difference between the expanded cell and the primary cell measures. If my hypothesis that the primary cell is the minimal possible container for the localization economies it true, the coefficient on the residual agglomeration should be insignificant. Otherwise, this might suggest that agglomeration expands beyond the primary three-digit industry – raion level cell into at least one dimension, industrial or territorial.

5.3. Econometric Model

I will follow a simplified form of Rosenthal and Strange's (2004) approach and directly estimate the production function in the same manner as Henderson (2003) did. Henderson estimated a standard production function of the firm *j* located in the area *m* in time *t*,

$$\ln(y_{jt}) = \boldsymbol{\alpha}\ln(\mathbf{X}_{jt}) + \boldsymbol{\beta}\ln\mathbf{E}_{mt} + \boldsymbol{\gamma}\ln\mathbf{I}_{jt} + \delta_t + \varphi_{mj} + \varepsilon_{mjt} \quad (5.1)$$

where y_{jt} is the output, $\mathbf{X}_{jt}$ is a vector of production inputs, E_{mt} is a vector of agglomeration variables, $\mathbf{I}_{jt}$ is a vector of institutional variables, δ_t is the time fixed effect; ϕ_{mj} is the three-digit industry and location fixed effect; and ε_{mjt} is the error term.

The unit of observation in equation (5.1) is an establishment, or a plant. Henderson (2003) estimated this function with the plant-level fixed effects model. He justified his method by noting that if more productive firms chose better location this difference would hopefully be captured by the firm fixed effect. I used a variation of this approach. In my preferred specification, I imply QMSA fixed effects and introduce an indicator variable for firms located in urban areas. I also assume a cluster-robust error structure in the residuals. The major reason for this diversion is the lack of variability in my localization measures over time and a different composition of the regions as in Henderson's sample. Unlike Henderson, who used five-year long periods over a fifteen year time span, my data are annual and form a five-year long panel. Preliminary results suggest that there is no sufficient time variation for estimation using firm level fixed effects.

Following Henderson (2003) and Nakamura (1985), I chose two agglomeration measures: the number of plants within a firm's cell and the employment within the same cell. First, I calculated the total number of firms and employment within a primary cell. Second, I calculated the total number of firms and employment within an extended cell. I considered two cases of the expansion of the primary cell. One was calculated at the level of a two-letter subsector within a raion, which represented the expansion into the industrial space, and another was calculated at the level of the three-digit industry within a QMSA, which was an expansion of the primary cell into the geographic space. In both primary cell and expanded cell measures, I then subtracted own employment and own plant from the total, respectively, to account only for the external effects. As a side note, in the case of industrial expansion the density of both plants and employment within the cell increased compared to the primary cell measure, and in the case of geographic expansion it went down. I need this notice to compare the results with Ciccone and Hall (1996). The third measure was the difference between the expanded cell measure and the primary cell measure. Algebraically, this may be shown as the following identity: $E_{Prim} + E_{Resid} = E_{Expand}$. This agglomeration measure should not necessarily be additive. Nakamura (1985) used an exponential function to measure agglomeration: $g(P) = \alpha P^{\alpha_P}$, where P was population in the city. Henderson (2003) used a similar approach. I also use the exponential functional form for my residual agglomeration measure:

$g(E) = \beta E_{Prim}^{\beta_{Prim}} E_{Resid}^{\beta_{Resid}}$. For the primary cell and the expanded cell, the functional form is similar.

I encountered a number of cases when the number of observations in the primary cells was not the same as for expanded cells or the residual measure. Also, this was not uncommon that the residual agglomeration was zero when only a single three-digit industry from the corresponding subsector was represented in the only raion or city within a QMSA. For such cases, the logarithmic function is undetermined, hence I chose $\ln(E_m + 1)$ transformation. I can justify this choice by noting that even if no external effect in the same industry-region cell or any residual agglomeration were present, a firm can still be operative, and this is not correct to exclude such stand-alone firms from the estimated sample.

I also experimented with other specifications. First, I modeled agglomeration for a firm as a linear combination between the agglomeration at the primary cell level and the agglomeration at the expanded cell level:

$$\beta_\delta \left(\delta E_{Prim} + (1-\delta)E_{Expand}\right).$$

I first estimated a (β_δ, δ) pair of coefficients within a non-linear model with a set of three-digit industry dummies, annual dummies and QMSA fixed effects (demeaning the data around QMSAs). This estimation strategy was not successful, as the δ coefficient tended to be out of (0, 1) range for most subsectoral equations both for plants and employment specifications. I conjectured that this indicated a possible propensity of agglomeration effects to group around one of the local extremum points of the $\delta E_{Prim} + (1-\delta)E_{Expand}$ segment. In other words, agglomeration effects are probably non-linear and even non-monotonous with respect to the distance between firms. Findings in the literature indirectly support this idea. Rosenthal and Strange (2003) have found that "localization effects attenuate rapidly in the first few miles, and slowly thereafter." Duranton and Overman (2005) found that even though localization takes place mostly between 0 and 50 km, the second "importance range" is the regional level, at 80-140 km, and not the immediate "sub-regional" level of 50-80 km.

I also fixed δ at several levels (such as 0.0001, 0.5 and 0.9999) and estimated β_δ by a usual fixed-effects model. To determine which specification should be preferred I ran the LR test comparing estimated models with different values of δ. This approach also did not show consistent results. The combination of the same aggregation levels seemed to work differently in various industries. This method could be seen as a rough imitation of Rosenthal and Strange's (2003) approach of drawing concentric circles around each firm and measuring the level of agglomeration within such a circle, but it lacked precision in measuring the distances, both in the industrial and geographic dimensions. Therefore I adhered to the specification (5.1).

Given the considerations about a possible difference between the agglomeration effects in a "core" industry-region cell and "residual" cell, I estimated three models:

$$\ln(y_j) = \boldsymbol{\alpha} \ln(\mathbf{X}_j, \mathbf{I}_j) + \beta_1 \ln E_{Prim} + \delta + \varphi_{mj} + \varepsilon_{mj} ; \tag{5.2}$$

$$\ln(y_j) = \boldsymbol{\alpha} \ln(\mathbf{X}_j, \mathbf{I}_j) + \beta_2 \ln E_{Expand} + \delta + \varphi_{mj} + \varepsilon_{mj} ; \tag{5.3}$$

$$\ln(y_j) = \boldsymbol{\alpha} \ln(\mathbf{X}_j, \mathbf{I}_j) + \beta_1 \ln E_{Prim} + \beta_3 \ln E_{Resid} + \delta + \varphi_{mj} + \varepsilon_{mj} . \tag{5.4}$$

If my hypothesis that the agglomeration effect is most pronounced at the level of the primary cell, is true then β_3 should be insignificant. If it is not the case, then the agglomeration effects in the given subsector spill beyond the primary cell within a geographical or industrial dimension. I estimated these three models separately for each two-letters subsector and for two sets of agglomeration measures, the plant count and the employment within the plant's cell. I used the QMSA fixed effects with subsidiary status of the firm and urban territory status as institutional dummies, a set of three-digit industry dummies, and annual dummies.

5.4. Estimation Results

5.4.1 General Results

For every subsector, all specifications resulted in approximately the same coefficients for the production factor elasticities. In all cases, the coefficients on capital and labor had expected signs and were highly significant. The differences in estimated values between all three specifications do not exceed one percent. That is why I present only the estimation results for the model (5.2). The estimates for the case when the agglomeration measure is the number of plants in the firm's industry-region cell are shown in Table 5.1, whereas Table 5.2 shows the estimates for the case when the agglomeration variable is employment in the firm's cell. The number of observation is the same for all three models. The models show a good fit to the data.

Henderson's (2003) paper is the closest to my work from the methodology standpoint. Henderson used only two industry groups, machinery and high-tech, which are roughly similar to Ukrainian subsectors DK and DL, respectively. He used the number of plants as the agglomeration measure. In both sectors, my coefficients are slightly greater: 0.047 vs. 0.033 in machinery and 0.103 vs. 0.061 in the "affiliate" sample and 0.038 in the "non-affiliate" sample in the high tech. Since he used materials and I did not, the coefficients on labor in my study are approximately twice as large compared to Henderson's. However, I may compare the sum of the production factor coefficients which indicate the returns to scale. In Henderson's study, the returns to scale are slightly below one in both industry groups (between .85 and .95), and he claims that the returns to scale are constant. In my study, I witness the increasing returns to scale, since the sums of the coefficients are slightly above unity in both sectors. The difference may be attributed to variations in methodology of measuring the capital and labor and to variation in the estimated model (firm fixed effects in Henderson (2003) vs. QMSA fixed effects in my study), but in general I notice a similarity in the results.

Another benchmark study is Brown et al. (2006) who used a similar data set (but an earlier and longer period) to explore privatization effects in Ukraine. They also separately estimated a firm-level production functions with firm fixed effects for ten subsectors. Unlike my results, in two subsectors out of ten in their study, the coefficients on capital turned out to be insignificant, whereas labor coefficients were also significant in all equations. In general, the capital marginal products in my study are found to be lower than in Brown et al. (2006), whereas labor elasticities were significantly lower. In my study, only four subsectors out of eleven have shown labor elasticities above one, and the rest were above 0.92, whereas in their study, the labor elasticity are above one in all subsectors. On the other hand, Brown et al. (2006) used firm-level data, whereas I used plant-level data. It is possible that aggregation of all firm's plants located in various regions into a single observation could be responsible for such a deviation. Another possibility for the discrepancy is that the data they used included a period when Ukraine underwent a hyper-inflation, and they had to use adjustment coefficients for capital. If those coefficients were calculated with an error, this could inflate their estimates of capital. My period of observation was free of high inflation, and I did not use any adjustments to the data. Despite all discrepancies in the observation period and estimation methods, I claim that coefficients on labor and capital in my model are reasonably close to those found in Brown et al. (2006).

In many ways my results are comparable to other agglomeration studies. Henderson (2003) measured localization economies at the level of an industry group in a county. He found that doubling the number of own-industry plants in the same county will increase firms' productivity by 10% to 13.5% in the high-tech and will have almost no effect in machinery. I will compare these results with the second column in the Table 5.3 which shows the estimation results of the model (5.3) for the same type of the agglomeration measure. We can see that my result is essentially the same for the subsector DL, which is close to Henderson's definition of the high tech industry group, whereas for the machinery (subsector DK) I also found a significant agglomeration effect at 9.8%. Nakamura (1985) used aggregated data and estimated localization and urbanization effects in the same model. He measured agglomeration at the level of a two-digit industry in a prefecture and in the city. If we compare my results (also the column two from the Table 5.3) with Nakamura's we will see that my result indicates slightly greater values. A possible reason for this is that I followed Henderson's (2003) approach and did not directly account for urbanization economies in the same model. That is why it is possible that my results may overestimate localization effects if a part of the effect attributes to general diversity of industries within a city and not only to the co-location of other firms in the same industry – region cell.

I may conclude that compared to other papers on agglomeration given the data, agglomeration measures and methodology used, my study indicates a presence of agglomeration economies in the Ukrainian manufacturing industries. The exact value of the effect is difficult to measure and compare precisely, but the estimates are generally in line with those found in other studies.

This is good news given a transition nature of the Ukrainian economy. One of possible counterarguments to the presence of localization effects in Ukraine and any other post-Soviet economy is a fact that under a planning system, location decision of firms should have reflected much broader markets, fixed prices and subsidized transportation costs. The cooperation between firms was not a result of their free choice, but determined by the central planners. Firms had little discretion in their choice of partners and suppliers. If any supply chain relationships within industries took place, they were not necessarily bound to a certain area and could even cross boundaries of republics. On the other hand, a system of compulsory placement of young graduated specialists to their first jobs (known as “raspredelenie,” or “allocation”, see Granick (1987), p.17), lack of performance-based compensation structure (Ericson (1991)) and an institute of mandatory residency registration (“propiska”) led to rigid job markets and limited possibilities for labor market pooling. This system essentially left out only one of the three traditional channels of Marshalian externalities, knowledge spillovers. I can conjecture that localization effects in the Soviet economy were probably smaller than in the Western countries, and urbanization effects were probably bigger if measured formally. Nevertheless I am not aware of any agglomeration studies in the Soviet Union. Massive privatization that took place mostly in nineties has set factor prices free and provided opportunities for firms to reallocate to a different sector. Even though by the beginning of 2001 privatization was almost over, my data show that about a quarter of firms have changed their manufacturing profile over next five years.

Nevertheless, as my study shows, even if these facts mattered before, after the massive privatization in nineties, the location of new firms and operation of formerly state firms was subjected to general economic laws. Firms apparently took into account the location of other firms in the industry. Then either new firms made a decision to start operations in the same place, or former state firms redirected their resources to the most productive use. In other words, firms in Ukraine seem to respond to the same principles as in every other market economy. I have found that localization economies are present in all sectors, and the values of the effect are comparable to findings from other Western studies.

5.4.2 Agglomeration Effects at the Level of the Subsector in the Raion

Now I turn to the differences between localization effects measured at different aggregation levels. In my first experiment I will compare the effect of aggregating into the industrial dimension, when the data are aggregated at the level of a subsector in the raion. In Table 5.3 I present the estimation results of the models (5.2) – (5.4) where the agglomeration measure is the number of plants in the region-industry cell. Table 5.4 presents the same for the employment measure.

First of all, I notice that effects for the plants measure are greater in almost all but one case compared to the effects when employment is used as a measure. This may suggest that agglomeration forces are formed at the level of the plant ownership or management rather than employment. When firms' owners make a location decision, they seem to care more about other firms in vicinity, whatever small they would be, rather than about labor to fill vacancies in those firms. In other words, my work suggests that knowledge spillovers and supply chain relationships are more pronounced compared to labor market spillovers. Certainly, this result is too generalized and probably averaged out since I do not know anything about skill levels of workers.

I compare the estimates of the "decomposed expanded cell" model (5.4) with the estimates from the "primary cell" model (5.2) and "expanded cell" model (5.3). I claim that estimating agglomeration economies at the same level of aggregation for all industries is not correct. Agglomeration effects reveal themselves at various scales in different subsectors of manufacturing.

Once we explore the results in Table 5.3, where the agglomeration measure is the number of plants in the cell, we can distinguish three separate cases. First is the case when the localization effects measured at the smallest, primary cell (three-digit industry and raion), do not significantly change if the agglomeration variables are measured at the level of the expanded cell, a subsector in the raion. In this case, the estimates of the agglomeration in the primary industry-raion cell and the expanded cell (the first and the second lines for each subsector in Table 5.3) are not too different, though the coefficients for the expanded cell are slightly greater in value and generally less statistically significant. At the same time, the effect of the "residual agglomeration," the third line for each subsector, which is measured as the difference between the expanded cell agglomeration and the primary cell agglomeration, turns out to be negligible. The decomposition of the agglomeration suggests that localization effects in these subsectors reveal at the smallest possible level within the primary cell. Put differently, most transactions probably take place within the same three digit industry, and there are no significant spillovers between firms from other industries within the same subsector.

In the second case, localization effects measured for the primary cell and the expanded cell are both significant. The decomposition of the expanded measure into "primary" and the "residual" effects shows that both effects are significant. However, the "decomposed" primary cell effect (the first row in the third column for every subsector in Table 5.3) is usually smaller in value than the "pure" primary cell effect. This suggests that firms are more actively involved into relationships with other firms both within and beyond own three-digit industry in the region. Productivity of such firms will grow if there is an increase in the total number of firms in the same subsector, regardless of the particular three digit industry

In the third case, we observe insignificance of the primary cell estimate in the "decomposed" model (5.4) and strong significance of the "residual" agglomeration coefficient, which suggests that the entire localization effect is gained at the level of the "residual agglomeration." This result is strange because it implies more urbanization-economy type story than purely localization effects. Firms in this case seem to gain in the productivity from the growth of the firms' number outside of their own primary cell.

If we explore the estimation results for the same three models, but using the employment as the agglomeration variable, as shown in Table 5.4 we will also distinguish the same three major cases. However, only in half of the cases the results from both types of agglomeration variables coincide. We now compare the results of these estimations and attempt to explain the discrepancies.

First, the results for subsectors DJ (basic metals and fabricated metal products) and DM (transport equipment), we observe the first case in both sets of estimates. In these two sectors, the agglomeration effect seems to reveal itself at the level of the own three digit industry within the same raion, and further expansion into the industry dimension seems to have no effect. It is remarkable that these two industries have the largest average firm size. Apparently, these firms are large enough to combine most production processes inside, and their productivity growth depends mostly on the availability of other firms and employment in the same industry.

For subsectors DA (food products, beverages and tobacco), DE (pulp, paper and paper products; publishing and printing), and DI (other non-metallic mineral products), agglomeration effects are evident both at the very local level and in the related industries in the same subsector, and this result is also the stable for employment and number of establishment measures. For three other sectors, DB (textiles and textile products, by construction pooled with DC, leather and leather products), DK (machinery and equipment "not elsewhere classified"), and DL (electrical and optical equipment), this effect is pronounced for the employment measure only, whereas for the plant measure, we observed the third type of the effect, the importance of the residual agglomeration. This result suggests that labor pooling both within the own three-digit industry and in the related industries from the same subsector is important for the industry growth in all these sectors. On the other hand, a positive productivity effect in the latter three sectors (DB, DK, and DL) occurs only from increasing the number of firms outside of the own primary cell. One of possible explanations is that the total number of establishments within these subsectors is sufficiently large: the three subsectors constitute about a third of observations in my data, and all six subsectors cover approximately three-quarters of the sample. This is not a perfect measure, but this may indicate near-satiation of the market, where agglomeration effects are counterweighted by competition forces. Firms enjoy the general increase of the local labor force, both in the own and related industries, but they start to compete for additional employees with each other. That is why we see insignificance for the number of plants within the primary cell in the decomposed measure.

In the subsector DD (wood and wood products), both employment and number of plants measures suggest that firms gain most agglomeration effects from other industries outside of the own cell. This may suggest a vertical integration with non-manufacturing industries (for example, with furniture producers) that we do not observe.

In the rest of the subsectors, the agglomeration estimates based on the number of plants and employment are different in outcomes. In the subsector DG (chemicals, chemical products and man-made fibers, by sample construction combined with DF, coke, refined petroleum products and nuclear fuel), I found that the employment measure indicates agglomeration effects only within the own primary cell, whereas for the number of plants in the rest of the subsector seems to be more important that in the own primary cell. Such result may indicate that even though firms do benefit from interactions with each other within their own industry, they suffer if the related industries experience employment growth, since this may distract their current and potential employees.

In the subsector DN (manufacturing, not elsewhere classified), the results are opposite to the previous case: the employment measure indicate positive localization effects from labor spillovers both within and outside the own primary cell, whereas the plant count measure indicates only within primary cell spillovers. This subsector includes service-like industries, as manufacture of furniture, jewelry, musical instruments, sports goods, games, toy, etc. This may mean that even though firms benefit from co-locating with other firms in the same primary cell, increasing employment in related industries is also important for them, suggesting rather an urbanization story than purely localization effects.

Finally, I cannot draw pervasive conclusion about the subsector DH (rubber and plastic products). On the one hand, the employment measure indicates very clear local effects as described in the first case above. On the other hand, decomposition of the expanded agglomeration measure based on the plant counts, resulted in both coefficients statistically insignificant. This is the second smallest subsector, which covers only about 3.7% of all observations in the sample, and the total number of observations may be insufficient for further conclusions.

Even though I used a different model than Moomaw (1998), I may draw a similar general conclusion, that “estimates of localization … economies using two-digit data are similar to estimates using three-digit data.” However, also in line with his findings, I should warn against using the same level of industrial aggregation for all industries alike. The level of aggregation which is appropriate for one industry may be entirely wrong for another. Agglomeration economies measured at the level of three-digit industry aggregation seem to exist and have the expected sign. Nonetheless, for several industries, localization effects spill over into related industries, which should also be accounted for.

5.4.3 Measuring Agglomeration Effects at the Level of Three Digit Industry in the QMSA

In my second experiment I compare agglomeration effects at the level of the primary cell and expanding this cell into the geographic dimension. Unlike in the previous section, this operation will decrease the density of firms in the cell, because in most cases, I will add raions that do not have economic activity in the given three-digit industry. Nakamura (1985) ran a similar experiment, when he compared the agglomeration effect measured at the level of a prefecture in Japan and at the level of cities. He found that expanding the region size level has increased the value of the localization coefficients in most industries, but in several cases either the more agglomerated effect fell, or turned insignificant, or both.

I present my results in Table 5.5 where the agglomeration measure is the number of establishments in the cell, and in Table 5.6 where the agglomeration is measure is the employment. In contrast to the industrial level results, the results are more homogenous across industries. First, when the agglomeration measure is the number of plants aggregating the data geographically in all cases but two renders insignificant coefficients. For those two subsectors when the results are significant, the value of the coefficient drops (subsector DI) or remains almost the same (subsector DK). Second, when I decompose agglomeration effects into those attributed to the primary industry-region cell and the "residual" agglomeration, I can clearly distinguish only two major cases. In the first case, the coefficient by the "own primary" component of the decomposed measure (the first row in the third column for each subsector in Table 5.5) is statistically significant and only slightly different in value from the own primary measure. In the second case, the coefficient by the "own primary" component of the decomposed measure is statistically insignificant, and the coefficient by the residual agglomeration measure is strongly statistically significant, but negative. There is one border case, when both coefficients in the decomposed measure are significant, but the coefficient by the residual agglomeration is still negative.

When the agglomeration measure is employment the localization effects are lower in value for the primary cell (the first column in Table 5.6) compared with the number of establishments case, do not substantially change (if significant) when agglomeration is measured at more aggregated geographical level (the second column), whereas the number of sectors where this coefficient is statistically significant is six rather than two. When I decompose the agglomeration effect into "own primary" cell and the residual, the value of coefficient by the "primary cell" component is very close to the primary cell measure and always significant, and the coefficient by the residual agglomeration component is insignificant in all cases but one, where it is not only significant, but also negative.

Hence, for the most subsectors (DA, DE, DG, DI, DJ, DK, DM, and DN), the estimates of localization economies coincide regardless the agglomeration measure. Aggregation of agglomeration effects at the higher level in the geographic hierarchy is wrong, since the agglomeration coefficients turn out to be insignificant. This finding is in line with the earlier works by Rosenthal and Strange (2003) and Duranton and Overman (2005) who found that agglomeration effects tend to attenuate quickly with the distance, and already after thirty to fifty kilometers the effects are not substantial. Since the average area of a raion in Ukraine is about 1000 square kilometers, expanding this distance clearly leads to insignificance of localization effects.

For the rest four industries, DB (textiles and textile products, by construction pooled with DC, leather and leather products), DD (wood and wood products), DH (rubber and plastic products), and DL ((electrical and optical equipment) expanding the geographic scale leads to disagglomeration indicated by at least one of the measures, either employment or establishments count. The negative and significant coefficient by the residual agglomeration variable suggests that if firms in these industries are engaged in interactions with firms in the same industry, these relationships are extremely local, and increasing the number of firms or employment outside of the own raion may be perceived as a negative factor for own productivity growth. A possible explanation to this phenomenon is the competition either for a market share, as in the case of the plants count, or for professionals in the same area, as in the case of employment.

I may also conclude that the localization effects reveal themselves at very geographically local levels and tend to attenuate with distance. Aggregating the data to large geographic regions such as MSA, states, or lands, does not add precision to the estimates. When the coefficients on more aggregated measures are significant, they are generally lower than those measured at local levels. Insignificance of the coefficients by the aggregated measures may be misinterpreted as the lack of agglomeration effects and result from the inference with urbanization effects in the same area.

5.4.4 Measuring Agglomeration Effects at the Level of Three Digit Industry in the Nearest Neighborhood

In my final experiment I also measure agglomeration economies at the level of three digit industries, but instead of a QMSA area I re-defined the geographic neighborhood for each firm. Using an adjacency matrix I created a "neighborhood area" that consists of raions which are adjacent to the raion of the firm. Such rearrangement of a firm's vicinity may be preferable over QMSA, when I implicitly implied that agglomeration effects are the same for all firms within a QMSA. QMSAs could include many raions, not all necessarily adjacent to each other, and stretch by dozens of kilometers. Therefore firms located in two adjacent raions within a QMSA were treated in the same way and assumed to have the same effect on each other as firms located in two raions at different ends of the QMSA. This seems to contradict the major finding from micro-geographic studies, such as Rosenthal and Strange (2003) or Duranton and Overman (2005) that agglomeration effects tend to attenuate with distance. To adjust for the difference in distances between firms within a QMSA one can measure agglomeration effects within approximately the same distance from a firm by constructing the neighborhood areas as described above. Agglomeration effects measured in such areas are methodologically similar to the spatial lags as described, for example, in Anselin (1988).

I restricted the sample only for the firms located within QMSAs to preserve consistency with my earlier methodology and findings, and constructed the neighborhood areas for all such firms. It is evident that not all neighborhood areas are contained entirely within the QMSA. For every neighborhood area I calculated the number of plants and total employment. Therefore the industry-region cell for each firm now includes all firms from the same three-digit industry located in the same "neighborhood" area.

I estimation result of models (5.2) through (5.4) are presented in Table 5.7 when agglomeration effects are measured as the number of plants in the same industry-region cell, and in Table 5.8 when the agglomeration measure is employment.

First I compare the estimation results between models (5.2) and (5.3), using the plant counts as the agglomeration effects measure. Aggregating the data into neighborhood areas does not change the significance of the coefficients compared to the case when the data are aggregated at the raion level, but the magnitude of the effects seems to significantly increase in all subsectors. The response in the external scale economies to doubling the number of plants within a neighborhood area is, on average, by three percent higher than in the case when the industry-region cell is a three-digit industry in the raion. Disaggregating the effects into the "core" cell and "residual" cell and estimating the model (5.4) indicates that in all sectors but four (DD: Wood and wood products, DI: Other non-metallic mineral products, DM: Transport equipment, and DN: Other manufacturing), the residual component adds virtually nothing to the total effect. The coefficient on the core cell component is essentially the same as in the estimation results for the model (5.2), whereas the residual component is insignificant and, in one case, even negative. The difference in the magnitude of the coefficients on the "core" component is also very small. This indicates that adding a "spatial lag" to the agglomeration measures and estimating the effect in the neighborhood does not significantly change the estimates, which supports the major finding from microgeographic studies about attenuation of agglomeration effects. For three other machinery-related industries (DI, DM, and DN), the residual component is significant and positive which may suggest reliance of firms in the same industries on availability of other firms in the neighboring raions and importance of relationships with them. In other words, measuring agglomeration effects at the most local level for these three industries is probably not correct, as their "significance" area is wider. The case of the industry DD (Wood and wood products) is an outlier in our results, since disaggregating the agglomeration effects into the core and the residual components leads to statistical insignificance of both coefficients, whereas estimation of the aggregated model (5.3) and the core model (5.2) shows the significance of the results. This may suggest some vertical relationships within the industry which are not immediately evident from the data.

When the agglomeration measure is the employment in the industry-region cell (Table 5.8), the conclusion about the aggregated data is the same as above: the coefficients in the aggregated model (5.3) are significant in all sectors but one (DH:

Rubber and plastic), and only slightly greater than in the core model (5.2). The average difference between the coefficients is less than one percentage point. Disaggregating the data (and estimating the model (5.4)) generally leads to the same conclusion as in the case of the plant counts: in all industries but three (DA: Food, beverages and tobacco, DI: Other non-metallic mineral products, and DK: Machinery and equipment n.e.c.), the agglomeration effects is revealed only at the core level, and the residual agglomeration does not add anything. This finding again supports the major conclusion from the micro-geographic studies about quick attenuation of the localization effects with an average distance between firms. As for the latter three industries, DA, DK, and DI, the reliance on employment in the neighboring raions may suggest positive effect of the labor pooling. In contrast with QMSA results (Table 5.6), when the residual agglomeration was always insignificant for the employment measure, measuring agglomeration in the neighborhood indicates positive spillovers between firms and suggests that using disaggregated data does not reveal the entire picture.

On the other hand, for the majority of industries, the localization economies seem to be very local and contained entirely within a firm's own raion.

5.5. Conclusion

In this paper I compared estimates of localization economies measured at various levels of aggregation in the industrial and geographic space. Following the recent research based on micro-geographic data, agglomeration effects are predicted to be local and tend to attenuate with physical distance. I decomposed an aggregated agglomeration measures into the local agglomeration ("primary cell") and the residual agglomeration. I found that measuring agglomeration variables at too aggregated levels of geographic hierarchy may lead to underestimates of localization effects or wrongfully show complete lack thereof.

When the data are aggregated in the industrial space, my findings are not that conclusive. It appears that optimal industry size where localization effects reveal themselves varies by industries. While firms in some industries interact with other firms mostly from the same three-digit industry group, in other industries firms also need to have contacts with firms from related industries within the same industry group. At the same time, in developed industries where competition is high, the growth in own productivity may be achieved at the cost of decrease in the number of firms or employment in the related industries. Such residual agglomeration effect may suggest competition for the market share or labor resources, or more complicated relationships within an industry.

In the future work, I plan to explore urbanization effects as well. Once a firm's owners make a decision about location in a particular place, they take into account the entire range of factors, both within the same industry and more general, such as the market potential, availability skilled and unskilled labor and other production factors in the area, the level of local business services and the like. From this point of view it is not entirely correct to limit the study only to a narrow range of localization effects that influence the firm. My study shows that other factors may be present which I did not account for. Including such factors would be a logical step foreword.

Another important contribution of this work was to show that different industries should not be treated in the same manner when we estimate agglomeration effects. Agglomeration effects reveal themselves differently for each firm. Averaging out these effects within an industry was inevitable at earlier stages of the research in the area, but far from being optimal. Similar to microgeographic studies where accounting for exact physical distances between firms has lead to unexpected conclusions about rapid attenuation of agglomeration in the physical space, there is a need to estimate similar distances in the industrial space. Probably, national industry classifications used for grouping industries into greater clusters are not the best way for measuring such distances, and more advanced approaches are called for the future line of research.

Tables

Table 5.1, Production Factors Elasticities for Primary Cell Specification, Agglomeration Variable: Number of Plants

Subsector	ln(Capital)	ln(Employment)	R-Square	N
DA: Food, beverages and tobacco	0.120***	1.190***	0.72	27,475
DB: Textiles, leather	0.146***	0.925***	0.64	13,211
DD: Wood and wood products	0.113***	1.111***	0.61	10,893
DE: Pulp, paper, publishing	0.210***	0.927***	0.63	20,616
DG: Chemicals, refined petroleum, coke	0.132***	0.960***	0.64	6,267
DH: Rubber and plastic	0.176***	1.064***	0.64	6,023
DI: Other non-metallic mineral products	0.097***	1.095***	0.7	10,274
DJ: Basic metals and fabricated metal	0.153***	0.962***	0.67	12,061
DK: Machinery and equipment n.e.c.	0.047**	0.971***	0.64	17,049
DL: Electrical and optical equipment	0.103***	0.927***	0.61	15,605
DM: Transport equipment	0.112***	0.930***	0.7	3,911
DN: Other	0.108***	0.998***	0.61	10,729

* p<0.1, ** p<0.05, *** p<0.01

Table 5.2, Production Factors Elasticities for Primary Cell Specification, Agglomeration Variable: Employment

Subsector	ln(Capital)	ln(Employment)	R-Square	N
DA: Food, beverages and tobacco	0.120***	1.191***	0.72	27,475
DB: Textiles, leather	0.148***	0.925***	0.64	13,211
DD: Wood and wood products	0.113***	1.111***	0.61	10,893
DE: Pulp, paper, publishing	0.211***	0.927***	0.63	20,616
DG: Chemicals, refined petroleum, coke	0.131***	0.957***	0.64	6,267
DH: Rubber and plastic	0.177***	1.066***	0.64	6,023
DI: Other non-metallic mineral products	0.094***	1.095***	0.7	10,274
DJ: Basic metals and fabricated metal	0.153***	0.961***	0.67	12,061
DK: Machinery and equipment n.e.c.	0.048**	0.971***	0.64	17,049
DL: Electrical and optical equipment	0.104***	0.926***	0.61	15,605
DM: Transport equipment	0.114***	0.930***	0.7	3,911
DN: Other	0.107***	0.994***	0.61	10,729

* p<0.1, ** p<0.05, *** p<0.01

Table 5.3, Agglomeration Measures: Number of Plants in the Same Subsector and the Same Raion

	Agglomeration variables	Model 1: KV3/Raion	Model 2: PSEK / Raion	Model 3: KV3 / Raion and PSEK / Raion Residual
DA	ln(# plants, KV3/Raion)	0.117***		0.048*
	ln(# plants, PSEK/Raion)		0.150***	
	ln(# plants, PSEK/Raion residual)			0.110***
DB	ln(# plants, KV3/Raion)	0.111***		0.031
	ln(# plants, PSEK/Raion)		0.171***	
	ln(# plants, PSEK/Raion residual)			0.137***
DD	ln(# plants, KV3/Raion)	0.045*		0.008
	ln(# plants, PSEK/Raion)		0.073***	
	ln(# plants, PSEK/Raion residual)			0.067***
DE	ln(# plants, KV3/Raion)	0.103***		0.053**
	ln(# plants, PSEK/Raion)		0.121***	
	ln(# plants, PSEK/Raion residual)			0.063***
DG	ln(# plants, KV3/Raion)	0.122**		0.043
	ln(# plants, PSEK/Raion)		0.137***	
	ln(# plants, PSEK/Raion residual)			0.107***
DH	ln(# plants, KV3/Raion)	0.065**		0.040
	ln(# plants, PSEK/Raion)		0.083**	
	ln(# plants, PSEK/Raion residual)			0.048
DI	ln(# plants, KV3/Raion)	0.146***		0.114***
	ln(# plants, PSEK/Raion)		0.144***	
	ln(# plants, PSEK/Raion residual)			0.074***
DJ	ln(# plants, KV3/Raion)	0.094***		0.061**
	ln(# plants, PSEK/Raion)		0.106***	
	ln(# plants, PSEK/Raion residual)			0.052
DK	ln(# plants, KV3/Raion)	0.068**		0.018
	ln(# plants, PSEK/Raion)		0.098***	
	ln(# plants, PSEK/Raion residual)			0.072***
DL	ln(# plants, KV3/Raion)	0.069**		0.007
	ln(# plants, PSEK/Raion)		0.103***	
	ln(# plants, PSEK/Raion residual)			0.095**
DM	ln(# plants, KV3/Raion)	0.110**		0.106**
	ln(# plants, PSEK/Raion)		0.100	
	ln(# plants, PSEK/Raion residual)			0.030
DN	ln(# plants, KV3/Raion)	0.144***		0.134***
	ln(# plants, PSEK/Raion)		0.118***	
	ln(# plants, PSEK/Raion residual)			0.016

Note: Sector descriptions: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and wood products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic metals and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Table 5.4, Agglomeration Measures: Employment in the Same Subsector and the Same Raion

	Agglomeration variables	Model 1: KV3/Raion	Model 2: PSEK / Raion	Model 3: KV3 / Raion and PSEK / Raion Residual
DA	ln(Employment, KV3/Raion)	0.056***		0.044***
	ln(Employment, PSEK/Raion)		0.089***	
	ln(Employment, PSEK/Raion residual)			0.047***
DB	ln(Employment, KV3/Raion)	0.088***		0.074***
	ln(Employment, PSEK/Raion)		0.091***	
	ln(Employment, PSEK/Raion residual)			0.037**
DD	ln(Employment, KV3/Raion)	0.032**		0.016
	ln(Employment, PSEK/Raion)		0.057***	
	ln(Employment, PSEK/Raion residual)			0.040***
DE	ln(Employment, KV3/Raion)	0.078***		0.053***
	ln(Employment, PSEK/Raion)		0.100***	
	ln(Employment, PSEK/Raion residual)			0.048***
DG	ln(Employment, KV3/Raion)	0.039***		0.034**
	ln(Employment, PSEK/Raion)		0.044**	
	ln(Employment, PSEK/Raion residual)			0.022
DH	ln(Employment, KV3/Raion)	0.057**		0.058**
	ln(Employment, PSEK/Raion)		0.057**	
	ln(Employment, PSEK/Raion residual)			-0.001
DI	ln(Employment, KV3/Raion)	0.049***		0.042***
	ln(Employment, PSEK/Raion)		0.063***	
	ln(Employment, PSEK/Raion residual)			0.034***
DJ	ln(Employment, KV3/Raion)	0.056***		0.049***
	ln(Employment, PSEK/Raion)		0.065***	
	ln(Employment, PSEK/Raion residual)			0.027*
DK	ln(Employment, KV3/Raion)	0.044***		0.032***
	ln(Employment, PSEK/Raion)		0.067***	
	ln(Employment, PSEK/Raion residual)			0.040***
DL	ln(Employment, KV3/Raion)	0.041***		0.032***
	ln(Employment, PSEK/Raion)		0.058***	
	ln(Employment, PSEK/Raion residual)			0.033*
DM	ln(Employment, KV3/Raion)	0.052***		0.050***
	ln(Employment, PSEK/Raion)		0.054	
	ln(Employment, PSEK/Raion residual)			0.023
DN	ln(Employment, KV3/Raion)	0.049***		0.029**
	ln(Employment, PSEK/Raion)		0.074***	
	ln(Employment, PSEK/Raion residual)			0.065***

Note: Sectors: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Table 5.5, Agglomeration Measures: Number of Plants in the Same Three-Digit (KV3) Industry and the Same QMSA

	Agglomeration variables	Model 1: KV3/Raion	Model 2: KV3 / QMSA	Model 3: KV3 / Raion and KV3 / QMSA Residual
DA	ln(# plants, KV3/Raion)	0.117***		0.113***
	ln(# plants, KV3 / QMSA)		0.063	
	ln(# plants, KV3 / QMSA residual)			-0.025
DB	ln(# plants, KV3/Raion)	0.111***		0.095***
	ln(# plants, KV3 / QMSA)		0.042	
	ln(# plants, KV3 / QMSA residual)			-0.070**
DD	ln(# plants, KV3/Raion)	0.045*		0.035
	ln(# plants, KV3 / QMSA)		-0.068	
	ln(# plants, KV3 / QMSA residual)			-0.091***
DE	ln(# plants, KV3/Raion)	0.103***		0.088***
	ln(# plants, KV3 / QMSA)		-0.021	
	ln(# plants, KV3 / QMSA residual)			-0.039
DG	ln(# plants, KV3/Raion)	0.122**		0.127**
	ln(# plants, KV3 / QMSA)		0.074	
	ln(# plants, KV3 / QMSA residual)			0.017
DH	ln(# plants, KV3/Raion)	0.065**		0.028
	ln(# plants, KV3 / QMSA)		-0.117	
	ln(# plants, KV3 / QMSA residual)			-0.109*
DI	ln(# plants, KV3/Raion)	0.146***		0.146***
	ln(# plants, KV3 / QMSA)		0.111***	
	ln(# plants, KV3 / QMSA residual)			-0.003
DJ	ln(# plants, KV3/Raion)	0.094***		0.092***
	ln(# plants, KV3 / QMSA)		0.084*	
	ln(# plants, KV3 / QMSA residual)			-0.009
DK	ln(# plants, KV3/Raion)	0.068**		0.072**
	ln(# plants, KV3 / QMSA)		0.066**	
	ln(# plants, KV3 / QMSA residual)			0.018
DL	ln(# plants, KV3/Raion)	0.069**		0.050
	ln(# plants, KV3 / QMSA)		0.011	
	ln(# plants, KV3 / QMSA residual)			-0.075***
DM	ln(# plants, KV3/Raion)	0.110**		0.100**
	ln(# plants, KV3 / QMSA)		0.120**	
	ln(# plants, KV3 / QMSA residual)			0.065
DN	ln(# plants, KV3/Raion)	0.144***		0.145***
	ln(# plants, KV3 / QMSA)		0.090	
	ln(# plants, KV3 / QMSA residual)			0.006

Note: Sectors: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Table 5.6, Agglomeration Measures: Employment in the Same Three-Digit (KV3) Industry and the Same QMSA

	Agglomeration variables	Model 1: KV3/Raion	Model 2: KV3 / QMSA	Model 3: KV3 / Raion and KV3 / QMSA Residual
DA	ln(Employment, KV3/Raion)	0.056***		0.055***
	ln(Employment, KV3 / QMSA)		0.058***	
	ln(Employment, KV3 / QMSA residual)			-0.010
DB	ln(Employment, KV3/Raion)	0.088***		0.086***
	ln(Employment, KV3 / QMSA)		0.074***	
	ln(Employment, KV3 / QMSA residual)			-0.017
DD	ln(Employment, KV3/Raion)	0.032**		0.032**
	ln(Employment, KV3 / QMSA)		0.014	
	ln(Employment, KV3 / QMSA residual)			-0.008
DE	ln(Employment, KV3/Raion)	0.078***		0.075***
	ln(Employment, KV3 / QMSA)		0.008	
	ln(Employment, KV3 / QMSA residual)			-0.011
DG	ln(Employment, KV3/Raion)	0.039***		0.039***
	ln(Employment, KV3 / QMSA)		0.023*	
	ln(Employment, KV3 / QMSA residual)			-0.014
DH	ln(Employment, KV3/Raion)	0.057**		0.061**
	ln(Employment, KV3 / QMSA)		0.001	
	ln(Employment, KV3 / QMSA residual)			0.014
DI	ln(Employment, KV3/Raion)	0.049***		0.048***
	ln(Employment, KV3 / QMSA)		0.047**	
	ln(Employment, KV3 / QMSA residual)			0.013
DJ	ln(Employment, KV3/Raion)	0.056***		0.056***
	ln(Employment, KV3 / QMSA)		0.039***	
	ln(Employment, KV3 / QMSA residual)			0.000
DK	ln(Employment, KV3/Raion)	0.044***		0.044***
	ln(Employment, KV3 / QMSA)		0.043***	
	ln(Employment, KV3 / QMSA residual)			0.010
DL	ln(Employment, KV3/Raion)	0.041***		0.037***
	ln(Employment, KV3 / QMSA)		0.013	
	ln(Employment, KV3 / QMSA residual)			-0.027**
DM	ln(Employment, KV3/Raion)	0.052***		0.050***
	ln(Employment, KV3 / QMSA)		0.030	
	ln(Employment, KV3 / QMSA residual)			0.022
DN	ln(Employment, KV3/Raion)	0.049***		0.049***
	ln(Employment, KV3 / QMSA)		0.022	
	ln(Employment, KV3 / QMSA residual)			-0.004

Note: Sectors: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Table 5.7, Agglomeration Measures: Number of Plants in the same Three Digit Industry and Immediate Raion Neighborhood

	Agglomeration variables	Model1: KV3 / Raion	Model2: KV3 / Neighborhood	Model3: KV3 / Neighb. Residual
DA	ln (# plants; KV3/Raion)	0.117***		0.109***
	ln (# plants; KV3/Neibourhood)		0.147***	
	ln (# plants; KV3/Neighb. Residual)			0.041
DB	ln (# plants; KV3/Raion)	0.111***		0.106***
	ln (# plants; KV3/Neibourhood)		0.173***	
	ln (# plants; KV3/Neighb. Resid.)			0.046
DD	ln (# plants; KV3/Raion)	0.045*		0.040
	ln (# plants; KV3/Neibourhood)		0.067***	
	ln (# plants; KV3/Neighb. Resid.)			0.021
DE	ln (# plants; KV3/Raion)	0.103***		0.103***
	ln (# plants; KV3/Neibourhood)		0.107***	
	ln (# plants; KV3/Neighb. Resid.)			-0.018
DG	ln (# plants; KV3/Raion)	0.122**		0.126**
	ln (# plants; KV3/Neibourhood)		0.141**	
	ln (# plants; KV3/Neighb. Resid.)			0.042
DH	ln (# plants; KV3/Raion)	0.065**		0.063*
	ln (# plants; KV3/Neibourhood)		0.079**	
	ln (# plants; KV3/Neighb. Resid.)			0.016
DI	ln (# plants; KV3/Raion)	0.146***		0.127***
	ln (# plants; KV3/Neibourhood)		0.187***	
	ln (# plants; KV3/Neighb. Resid.)			0.084***
DJ	ln (# plants; KV3/Raion)	0.094***		0.092***
	ln (# plants; KV3/Neibourhood)		0.127***	
	ln (# plants; KV3/Neighb. Resid.)			0.051*
DK	ln (# plants; KV3/Raion)	0.068**		0.063**
	ln (# plants; KV3/Neibourhood)		0.095***	
	ln (# plants; KV3/Neighb. Resid.)			0.050
DL	ln (# plants; KV3/Raion)	0.069**		0.070**
	ln (# plants; KV3/Neibourhood)		0.106***	
	ln (# plants; KV3/Neighb. Resid.)			0.016
DM	ln (# plants; KV3/Raion)	0.110**		0.117**
	ln (# plants; KV3/Neibourhood)		0.159***	
	ln (# plants; KV3/Neighb. Resid.)			0.162**
DN	ln (# plants; KV3/Raion)	0.144***		0.142***
	ln (# plants; KV3/Neibourhood)		0.169***	
	ln (# plants; KV3/Neighb. Resid.)			0.077**

Note: Sectors: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Table 5.8, Agglomeration Measures: Employment in the same Three Digit Industry and Immediate Raion Neighborhood

	Agglomeration variables	Model1: KV3 / Raion	Model2: KV3 / Neighborhood	Model3: KV3 / Neighb. Resid.
DA	ln (Empl.; KV3/Raion)	0.056***		0.053***
	ln (Empl.; KV3/Neibourhood)		0.083***	
	ln (Empl.; KV3/Neighb. Resid.)			0.023***
DB	ln (Empl.; KV3/Raion)	0.088***		0.086***
	ln (Empl.; KV3/Neibourhood)		0.104***	
	ln (Empl.; KV3/Neighb. Resid.)			0.019
DD	ln (Empl.; KV3/Raion)	0.032**		0.031**
	ln (Empl.; KV3/Neibourhood)		0.042***	
	ln (Empl.; KV3/Neighb. Resid.)			0.005
DE	ln (Empl.; KV3/Raion)	0.078***		0.080***
	ln (Empl.; KV3/Neibourhood)		0.077***	
	ln (Empl.; KV3/Neighb. Resid.)			-0.010
DG	ln (Empl.; KV3/Raion)	0.039***		0.040***
	ln (Empl.; KV3/Neibourhood)		0.032**	
	ln (Empl.; KV3/Neighb. Resid.)			-0.006
DH	ln (Empl.; KV3/Raion)	0.057**		0.057**
	ln (Empl.; KV3/Neibourhood)		0.046	
	ln (Empl.; KV3/Neighb. Resid.)			-0.001
DI	ln (Empl.; KV3/Raion)	0.049***		0.040***
	ln (Empl.; KV3/Neibourhood)		0.069***	
	ln (Empl.; KV3/Neighb. Resid.)			0.037***
DJ	ln (Empl.; KV3/Raion)	0.056***		0.056***
	ln (Empl.; KV3/Neibourhood)		0.057***	
	ln (Empl.; KV3/Neighb. Resid.)			0.006
DK	ln (Empl.; KV3/Raion)	0.044***		0.039***
	ln (Empl.; KV3/Neibourhood)		0.058***	
	ln (Empl.; KV3/Neighb. Resid.)			0.021**
DL	ln (Empl.; KV3/Raion)	0.041***		0.041***
	ln (Empl.; KV3/Neibourhood)		0.049***	
	ln (Empl.; KV3/Neighb. Resid.)			-0.001
DM	ln (Empl.; KV3/Raion)	0.052***		0.049***
	ln (Empl.; KV3/Neibourhood)		0.060***	
	ln (Empl.; KV3/Neighb. Resid.)			0.035
DN	ln (Empl.; KV3/Raion)	0.049***		0.047***
	ln (Empl.; KV3/Neibourhood)		0.053***	
	ln (Empl.; KV3/Neighb. Resid.)			0.012

Note: Sectors: DA: Food, beverages and tobacco; DB: Textiles, leather; DD: Wood and products; DE: Pulp, paper, publishing; DG: Chemicals, refined petroleum, coke; DH: Rubber and plastic; DI: Other non-metallic mineral products; DJ: Basic and fabricated metal; DK: Machinery and equipment n.e.c.; DL: Electrical and optical equipment; DM: Transport equipment; DN: Other manufacturing n.e.c.

Chapter 6. Conclusion

Using a rich panel of Ukrainian manufacturing establishments in 2001-2005, I analyzed patterns in agglomeration economies in Ukraine. An economic downturn of the nineties was followed by a decline both in the total population and employment in all sectors of the economy. On the other hand, our data indicate that a massive relocation of resources between regions and between sectors took place. First, employment seems to decrease in the rural areas, whereas some cities experienced growth in this factor. Second, economic activity seems to become more spatially concentrated over time. Inequality of spatial distribution of employment between regions in Ukraine appears to have increased, mostly at the account of urbanized areas. This finding is also partially supported by the analysis of the spatial correlations between regions. While cities and suburban raions enjoyed an increase in the total number of firms, rural areas revealed a reverse trend. This may suggest an involvement of suburban areas into the production processes of the cities.

Analysis of agglomeration indices also indicated an increased geographic concentration for a number of industries. At the same time, the structure of the economy is still close to one observed in the former Soviet Union. Even though the average size of firms has been decreasing in all years since the beginning of the transition period, large firms still dominate the economy, employing the absolute majority of the total manufacturing employment in Ukraine. Therefore, significant industrial concentration is observed, which may clutter the estimates of the geographic concentration. Yet it is remarkable that agglomeration patterns for several industries in Ukraine (with "Publishing" being the most striking example) resemble those in the Italian economy, as measured in Lafourcade and Mion (2007). This leads to a possible conclusion that Ukrainian economy has been adjusting to similar patterns observed in more developed countries.

In my analysis of external scale economies I first concentrated on two industry groups, machinery and higher technology sectors. My major goal was two-fold. First, I planned to compare the patterns in agglomeration economies with those found in Henderson (2003), and therefore I prepared the data to be most comparable to his and followed his estimation strategy. Second, I amended this comparison study with the analysis of the cross-effects between agglomeration economies and the ownership structure, as the share of state-owned and previously state-owned firms in Ukraine is significantly greater than in any developed economy. I found that localization economies are indeed present in both industry groups, which supports Henderson's (2003) findings, and that the effects are localized geographically. The major channel of localization effects seems to be a firm's management or relationships between owners, but not the employment or common labor pool in the area. I also found that majorly foreign owned firms seem to benefit the most from the localization economies and enjoy the effects of spillovers between firms to the greatest extent, as indicated by the productivity growth. Majorly domestic private firms also appear to get more effects from agglomeration compared to the state owned firms, especially in the high tech sector. This finding is important as it indicates that firms in the same industry and the same location may differ in the extent they can benefit from the same environment. The difference is likely to be attributed to the managerial structure of the firms and variation in the incentives of the owners and managers. This issue deserves further investigation.

I found as well that urbanization economies appear to be present in both industry groups and are also substantially localized. On the other hand, size-related measures as indicators of urbanization economies, though standard in the classic agglomeration literature, seem not to be the most adequate for studies when firm-level data are used. According to Jacobs (1969), urbanization economies are channeled not only through the city size, but rather through the level of diversity and innovative activities in the city. The size measures do not capture such relationships, and further research of the "diversity measures" is required to measure urbanization economies more precisely.

In the last chapter of my dissertation I tested a hypothesis that localization economies tend to attenuate with distance. This issue is important in light of the fact that sometimes, only aggregated data are available, whereas aggregation may lead to incorrect inference about agglomeration effects.

The "distance" should be understood not only in the geographic sense, but also as a difference between production processes, which is referred to as the "industrial distance". I decomposed the standard measures of agglomeration economies into the "core" and the "residual" effects and estimated three models separately for manufacturing subsectors: one for the smallest industry-region cell, one for the cell expanded into the geographic or the industrial dimension, and one for the decomposed effect. I found that localization economies are indeed localized geographically, and aggregating the data into larger geographic areas is not always justified, as the agglomeration coefficients in aggregated data tend to overestimate the external scale effects. This result is robust to different ways of data aggregation and generally support findings from micro-geographic studies that the agglomeration economies typically start to attenuate after thirty to fifty kilometers. However, aggregation of the data in the industrial space does not lead to the same conclusive outcomes. It appears that firms in certain industries seek for spillovers beyond their own industry group, but within the same (larger) sector. On the other hand, this effect appears to be sector-specific. Therefore aggregating the data into larger industry groups does not always lead to the same estimates of agglomeration economies, since the spillovers beyond own sector (so called "residual agglomeration") may play an important role for the firm's performance. The issues of the "optimal" industry size for agglomeration analysis and better ways of measuring distances between firms in the industrial space are promising venues of the future research.

Bibliography

1. Anselin, Luc, “Spatial econometrics: methods and models”, Boston: Kluwer Academic Publishers, 1988.
2. Black, Duncan and Vernon Henderson (2003), "Urban evolution in the USA", Journal of Economic Geography, 3:343-372.
3. Brown, David, John S. Earle and Almos Telegdy, 2006. "The Productivity Effects of Privatization: Longitudinal Estimates from Hungary, Romania, Russia, and Ukraine," Journal of Political Economy, University of Chicago Press, vol. 114(1), pages 61-99.
4. Brown, David and John S. Earle, "The Productivity Effects of Privatization in Ukraine: Estimates from Comprehensive Manufacturing Firm Panel Data, 1989-2005," Staff Working Papers 07-137, W.E. Upjohn Institute for Employment Research, 2007
5. Devereux, Michael P., Griffith, Rachel and Simpson, Helen (2004),"The geographic distribution of production activity in the UK," Regional Science and Urban Economics, vol. 34(5), pages 533-564
6. Dunaev E. P (n.d.), «Концентрация социалистического производства» [Concentration of Socialist Manufacturing Industries] , in “Bolshaya Sovetskaya Enciklopediya” [Big Soviet Encyclopedia], Vol. “K”, from http://bse.sci-lib.com/article064296.html
7. Duranton, Gilles and Diego Puga, 2001."Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," American Economic Review, American Economic Association, vol. 91(5), pages 1454-1477.
8. Duranton, Gilles and Henry G. Overman (2008), "Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data," Journal of Regional Science, vol. 48(1), pages 213-243
9. Dyker, David (1983), “The Process of Investment in the Soviet Union,” Cambridge, New York and Sydney: Cambridge University Press
10. Eberts, Randall W. and McMillen, Daniel P., 1999. "Agglomeration economies and urban public infrastructure," in: P. C. Cheshire and E. S. Mills (ed.), Handbook of Regional and Urban Economics, edition 1, volume 3, chapter 38, pages 1455-1495.
11. Ellison, Glenn and Glaeser, Edward L (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," Journal of Political Economy, vol. 105(5), pages 889-927.
12. Glaeser, Edward L., Hedi D. Kallal and Jose A. Scheinkman and Andrei Shleifer, 1992. "Growth in Cities," Journal of Political Economy, University of Chicago Press, vol. 100(6), pages 1126-52, December.
13. Henderson, Vernon, Kuncoro, Ari and Turner, Matt, 1995."Industrial Development in Cities," Journal of Political Economy, University of Chicago Press, vol. 103(5)
14. Henderson, Vernon, "Efficiency of Resources Usage and City Size," Journal of Urban Economics, 1986, 19, pp. 47-70.
15. Henderson, Vernon, “Marshall's Scale Economies” Journal of Urban Economics, 2003, 53, pp. 1-28
16. Holmes, Thomas J. and John J. Stevens, "Geographic Concentration and Establishment Scale," The Review of Economics and Statistics, 2002, vol. 84(4), pages 682-690, 07.
17. Jacobs, Jane (1969) “The Economy of Cities”, New York, Vintage Books

18. Lafourcade, Miren and Mion, Giordano, "Concentration, agglomeration and the size of plants," Regional Science and Urban Economics, 2007, vol. 37(1), pages 46-68.
19. Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, vol. 70(2), pages 317-341.
20. Lonsdale, Richard (1965) "The Soviet Concept of the Territorial-Production Complex," *Slavic Review*, Vol. 24, No. 3, pp. 466-478
21. Maré, David, "Concentration, Specialization and Agglomeration of firms in New Zealand," Working Papers 05_12, Motu Economic and Public Policy Research, 2005
22. Marshall, A. (1920) "*Principles of Economics"*, New York: Macmillian
23. Maurel, Francoise and Beatrice Sedillot, "A measure of the geographic concentration in french manufacturing industries," Regional Science and Urban Economics, 1999, vol. 29(5), pages 575-604.
24. Moomaw, Ronald L, 1981."Productivity and City Size? A Critique of the Evidence [Are There Returns to Scale in City Size?]." The Quarterly Journal of Economics, MIT Press, vol. 96(4), pages 675-88.
25. Moomaw, Ronald L., 1983. "Is population scale a worthless surrogate for business agglomeration economies?" Regional Science and Urban Economics, Elsevier, vol. 13(4), pages 525-545.
26. Nakamura, Riohei. (1985) "Agglomeration Economies in Urban Manufacturing Industries: a Case of Japanese Cities", *Journal of Urban Economics*, 17: 108–124
27. O'Donoghue, Dan and William Gleave (2004), "A Note on Methods for Measuring Industrial Agglomeration", *Regional Studies*, Vol. 38.4, pp. 419–427
28. Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, vol. 64(6), pp. 1263-97.
29. Paskhaver O., L. Verkhovodova, and N. Tereschenko, *Завершальний етап приватизації в Україні* [Finalizing stage of Privatization in Ukraine], О. Й. Пасхавер, Л. Т. Верховодова, Т. Є. Воронкова, Н. В. Терещенко. В надзаг.: Центр економічного розвитку. – Kyiv, 2003. – p. 11.
30. Quigley, John M, 1998."Urban Diversity and Economic Growth," Journal of Economic Perspectives, American Economic Association, vol. 12(2), pages 127-38.
31. Rosenthal, Stuart S. and William C. Strange, 2003. "Geography, Industrial Organization, and Agglomeration," The Review of Economics and Statistics, MIT Press, vol. 85(2), pages 377-393.
32. Rosenthal, Stuart S. and Strange, William C., 2004. "Evidence on the nature and sources of agglomeration economies," in: J. V. Henderson and J. F. Thisse (ed.), Handbook of Regional and Urban Economics, edition 1, volume 4, chapter 49, pages 2119-2171 Elsevier.
33. Segal, D. (1976) "Are There Returns to Scale in City Size?"*Review of Economics and Statistics*, 58: 339–450
34. State Committee of Statistics of Ukraine (Derzhkomstat), "Statystychnyy Schorichnyk Ukrajiny za 2005 rik" ["Statistical Yearbook of Ukraine, 2005"], Kyiv, 2006
35. Sveikauskas, Leo (1975) "The Productivity of Cities", *Quarterly Journal of Economics*, 89: 393–414

VITA

VOLODYMYR V. VAKHITOV

PERSONAL:

Birth Date: January 21, 1976
Birth Place: Barnaul, Russian Federation (former USSR)
Citizenship: Ukraine

EDUCATION

National University of Kyiv-Mohyla Academy:
M.A. Economics, 1999
B.A. Economics, 1997

EMPLOYMENT

Kyiv School of Economics, Kyiv, Ukraine
Research Associate and Data Specialist, August 2006 to December 2007
Visiting Instructor in Economics, September 2003 to May 2006
Research Assistant, January 2004 – August 2006

University of Kentucky, Lexington, KY, USA
Teaching Assistant / Primary Instructor, August 1999 to May 2003

PROFESSIONAL PUBLICATIONS

J. David Brown, John S. Earle and Volodymyr Vakhitov, "Wages, layoffs, and privatization: Evidence from Ukraine," *Journal of Comparative Economics*, Volume 34, Issue 2, June 2006, Pages 272-294.