European market integration, R&D agglomeration economies, and the location choice of firms in Poland

Astrid Krenz*, Agnieszka Gehringer**

Abstract

The establishment of a common European market, with the aim to eliminate internal barriers of different natures, implies an environment in which standard location determinants for firm activities might operate differently than usual or in which new location forces assume importance. Poland has had an exemplar experience as a formerly planned economy that managed the passage towards a market economy within a relatively short period of time. The country is characterized by comparatively low R&D expenditures per capita as well as few patent applications compared to other European Union countries in the past decade. Aside from the previous literature's focus on industrial and service agglomeration economies, we investigate the impact of agglomeration economies stemming from the R&D sector and innovative activities for the location choice of firms in Poland. We also assess which role the accession to the European Union played in this process. We use a panel data set comprised of Polish NUTS II regions over the period 2003 to 2010. The analysis differentiates the effects between domestic, foreign and entrepreneurial firms in the economy and enriches the former literature, which focused on the location decision of FDI. We find a strong impact of R&D agglomeration economies for the location choice of all firms, a positive impact of EU membership on the location decision of entrepreneurial firms, a positive impact of innovative activity for the location decision of foreign firms and a contingent effect of EU membership on the decision of foreign firms to locate closer to the German-Polish border.

Keywords: Agglomeration Economies, Border effects, Count Data, Eastern Europe, European Integration, R&D activities

JEL classification: C23, F15, F23, R11, R12

*Corresponding author: Astrid Krenz, University of Goettingen, Department of Economics, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany, astrid.krenz@wiwi.uni-goettingen.de, phone: 0049 551 397296.

**Agnieszka Gehringer, Flossbach von Storch Research Institute, Ottoplatz 1, 50226 Cologne, Germany, agnieszka.gehringer@fvsag.com, phone: 0049 221 3388284.

INTRODUCTION

The analysis of firms' location choices and of factors that contribute to geographic concentration of economic activity has attracted attention not only of scientists, but also of policy-makers. Starting with the early contributions of location theory, which paved the way to urban and regional economics, and more recently with the lively development in the field of New Economic Geography (NEG), authors investigated the determinants and consequences of spatial concentration. This interest is mainly due to important implications in terms of internal and international competitiveness of industries and services and the distribution of employment and income. Moreover, the establishment of a common European market, with the aim to eliminate internal barriers of different types, implies an environment in which standard location determinants operate differently than usual or in which new location forces assume importance.

The past empirical investigations of industrial location within Europe are dedicated mainly to the old member states: Take for example Figueiredo et al. (2002), who analyze the location determinants for Portugal, Arauzo-Carod and Manjón-Antolín (2004) do so for Spain, Crozet et al. (2004) for France, Disdier and Mayer (2004) for French multinationals and Bade and Nerlinger (2000) for Germany. Regarding new EU member states from Eastern Europe, to the best of our knowledge the only direct attempts to investigate location forces for Poland were made by Cieślik (2005a and 2005b) and Chidlow et al. (2009). The authors investigate the determinants of location only of foreign multinational enterprises (MNE) in Poland. Similarly, the study by Gauselman and Marek (2012) compares the factors determining the location choice of MNEs in 33 regions (NUTS II data) of East Germany, Poland and the Czech Republic, and Pusterla and Resmini (2007) investigate the location choices of foreign firms in Bulgaria, Hungary, Poland and Romania (using NUTS II data). Hilber and Voicu find that in Romania services agglomeration economies are a decisive factor for foreign firms' location choice and detect that the effects have been underestimated in previous studies. In a recent contribution, Vechiu and Makhlouf (2014) investigate specialization patterns of manufacturing industries for the EU27, including Central and Eastern European countries (CEEC). Their main result is that the CEEC first tended to diversify their production and finally specialized through the inflow of FDI. They explain that during the communist era, only a few large state-owned enterprises existed that did not meet foreign investors' request of efficiency, which resulted in large layoffs and consequently a decreasing degree of specialization (since specialization is measured by employment shares). The study reveals that low labor costs are a decisive factor for attracting FDI in CEEC.

Our contribution is aimed at filling the gap of inquiries into the geographical location of firm activity in CEEC, by analyzing the dynamics of location choices of all firms, taking into account the location decision of

domestic and foreign firms and entrepreneurs in Poland in the last decade. Poland has had an exemplar experience as a formerly planned economy that managed the passage towards a market economy within a relatively short period of time. The process of transition involved the establishment of a new economic and institutional framework with dynamically progressing privatization and industrialization and intense restructuring of enterprises (Carlin et al. 1995; Fidrmuc 2007; Robinson 2004). These intensive efforts to establish a market-based economic structure –which in many areas of transformation were made shock-therapeutically rather than gradually— were crucial in fulfilling economic and legal requirements deriving from EU-accession rules. During the entire transformation since 1989, Poland has become a main FDI receiving country in Eastern and Central Europe. Finally, in 2004, Poland became a member of the EU.

In contrast to the previous literature, we do not solely investigate the decisions to locate taken by foreign investors, but consider domestic activity and the activities of entrepreneurs, as well. This focus is particularly important from the point of view of European cohesion policy goals, which aim to improve regional indicators such as research and development performance, regional competitiveness and the business environment, the share of tertiary educational attainment, further sustainability and inclusiveness goals from which all firms in the economy will benefit.

In the previous literature, only agglomeration economies stemming from the industrial and less frequently from the services sector were investigated. This paper specifically investigates the importance of R&D agglomeration economies and the innovation activity for the firms' decision to locate. We believe this is an especially important aspect to consider for the case of Poland, since its expenditures on R&D per capita and its innovative activities as measured in terms of patent applications to the European Patent Office (EPO) are among the lowest in the European Union. Establishing significant and meaningful effects stemming from the generation and diffusion of knowledge in the Polish economy calls for an improved political support of the knowledge sector's development to control firm dynamics and in particular firm location decisions in Poland.

Using a negative binomial regression analysis, we find that agglomeration economies stemming from the industrial, services and R&D sector, human capital and infrastructure positively impact the location choice of firms in the Polish economy. There exist considerable differences between domestic, foreign and new firms in their location decisions across the regional economy. Specifically, accession to the European Union positively impacted the number of entrepreneurial firms and fostered the tendency of foreign firms to locate closer to the German-Polish border. Moreover, accession to the European Union increased the impact of innovative activity for the location decision of foreign firms. Furthermore, market demand and human capital appear to be significant, which contrasts with previous findings in the literature.

The paper is organized as follows. The next section reviews the theoretical background and findings of past contributions. We discuss agglomeration economies, labor market effects and differences between entrepreneurs and incumbents. Section 3 discusses the data and empirical methodology and presents and comments on the results obtained from a negative binomial regression analysis. The last section concludes.

THEORETICAL BACKGROUND

The Role of Agglomeration Economies for the Geographical Location of Firms

Alfred Marshall (1890) made one of the earliest contributions, pointing out the dynamics of geographic concentration between regions. His conceptual work identified a separate category of external economies of scale consisting of at least partially unpaid advantages to firms located in a certain region. Such advantages derive from the presence of inter-linkages in business activities with other firms in the same region. Such economies of scale are not internal, but external to a given firm. Subsequently, the importance of economic linkages between firms at the regional level has been extensively analyzed in the NEG field pioneered by Paul Krugman (1991).

More precisely, due to the proximity of upstream and downstream firms as well as of workers, each firm experiences a unit cost advantage in an analogous manner as in the case of internal economies of scale. There are at least three different sources of such unit cost reduction. First, from the supply side, locating close to suppliers will reduce input factor costs for producers, who save transportation and transaction costs. Analogously, from the demand side, locating close to demand will reduce transport costs and induce benefits of improved market access for producers. Second, proximity between suppliers of production factors (capital and labor) and their users makes it possible for the former to develop specialized technical properties and skills necessary for a more efficient production of the latter. Third, being located together leads to a more intensive exploitation of knowledge externalities, generated both at the demand- and at the supply-side.

Moreover, location benefits might refer to institutional advantages, related, for instance, to the quality of infrastructure. However, such positive effects may be only conditionally available. Indeed, apart from centripetal forces that attract firms to certain regions, there could be opposite centrifugal forces that will drive the firms out of the local center. Among centrifugal factors, the level and intensity of congestion or the relatively high unit input prices might exercise a negative influence on the geographic concentration (Krugman 1996).

Subsequent theoretical work in this area focused on more precise causes of agglomeration. Helsley and Strange (1990) demonstrated that agglomeration economies can arise from beneficial pooling of specialized

labor market forces. Goldstein and Gronberg (1984) focus on functionality advantages deriving from the opportunity to share the same supplier at a location. Glaeser (1999) identifies a particular role played by localized knowledge spillovers as promoting agglomeration dynamics.[1]

Against the well-developed theoretical basis underlying the spatial distribution of economic activities, empirical investigations of the location issue have a much shorter history. After the contribution of Marshall, only scant attention has been paid to verifying the hypotheses, with representative attempts by Carlton (1983), Bartik (1985) and Luger and Shetty (1985). Only after significant improvement was made through the emergence of NEG have new motivation and innovative empirical approaches enriched the bulk of evidence confirming the positive role played by agglomeration economies (Head et al. 1995; Guimarães et al. 2000; Rosenthal and Strange 2001; Cieślik 2005a, to name just a few).[2]

Agglomeration economies refer to different regional characteristics. Nevertheless, some attempts have been made to identify a single all-compressing measure that expresses the region's volume of economic activity. This approach encountered opposition of authors who argued that averaging different aspects of regional economic activity makes the measurement of agglomeration dynamics become imprecise (Head et al. 1995; Guimarães et al. 2000). Cieślik (2005a) addresses such concerns by distinguishing between four types of agglomeration economies: overall agglomeration economies, approximated by the regional GDP volume and measuring both the demand- and supply-side regional economic volume (Head and Ries 1996; Broadman and Sun 1997); urbanization economies measured by the percentage share of the population living in the urban areas and expressing the potential informational advantage of regions with high urban density (Glickman and Woodward 1988; Coughlin and Segev 2000)[3]; and industry and service specific economies, expressed by the respective shares of employment in the secondary/tertiary sector and representing the specialization advantage in a specific industrial or service sector (Woodward 1992; Smith and Florida 1994; Guimarães et al. 2000).

Labor Market Determinants

Among the standard determinants of location choice, agglomeration forces –as described above— belong to the group of factors that are prevalently taken into consideration. Additionally, other forces related to the labor market and to some other regional characteristics have been intensively investigated, especially in the context of the location decision connected with FDI inflows. Whereas agglomeration economies have often been confirmed to play a significant role in positively influencing regional attractiveness, evidence regarding the labor market forces –at least for the industrialized countries— was rather mixed. This more specifically concerns the characteristics of the labor force, its availability and educational attainment. Past estimations displayed either

that the higher the level of unit wages, the lower the attractiveness of the region for location purposes (Crozet et al. 2004) or no significant effects exist at all (Woodward 1992). Regarding the level of education, the literature demonstrates positive effects of human capital as proxied by the level of educational attainment among the local labor force (Alama-Sabater et al. 2011; Coughlin and Sergev 2000), although a negative relationship has been found as well (e.g., Arauzo-Carod and Viladecans-Marsal 2009). One explanation for the negative relationship might be that firms would have to pay higher wages if the labor force is better educated. In his study for the Polish economy, Cieślik (2005a) finds no significant effect for the level of schooling on the location choice of multinational enterprises. Arauzo-Carod (2013) finds evidence for a positive relationship between the level of educational attainment and firms' location decision when measuring human capital over a wider spatial area. Arguably, employers look for potential employees not only locally but also in surrounding areas. Similar arguments are provided by Holl (2004).

Further Aspects in Agglomeration Studies

The distinction between newly established and incumbent enterprises seems to be non-negligible. Indeed, the literature found a different pattern of location decisions between newly established and existing firms. In particular, the Marshallian forces were found to be less relevant for the location decision of new firms (Rosenthal and Strange 2001). The authors argue that the incumbent firms have managed to survive and thus cluster in an area of beneficial agglomeration economies. Different effects between start-up and relocating firms are also found by Holl (2004) and Manjón-Antólin and Arauzo-Carod (2011). Specifically, Manjón-Antólin and Arauzo-Carod find that location economies matter more for the start-up than for relocating firms.

EMPIRICAL SET-UP

Data Issues and Selection of Variables

For the analysis, we use data from the Statistical Regional Yearbooks of the Polish Central Statistical Office's online database. NUTS II data were retrieved, taking observations for the 16 Polish voivodeships for the years from 2002 to 2010. The time period was chosen due to data availability. From Eurostat's online database, data on the number of patents were extracted, as was a measure for R&D expenditures.[4] Gathering more disaggregated data, like NUTS III data, was not possible because of limited data availability. Several important variables, for example, measures for agglomeration economies from the industry, services and R&D sector, were only available at the NUTS II level. However, we have reasons for grounding our analysis on NUTS II data,

since the labor market with its mobile workers can be assumed to show its effects only at a higher level of aggregation (Arauzo-Carod 2013; Holl 2004). Moreover, the market demand can be assumed to be dispersed and to show effects for the location decision of firms only over a wider spatial area (Cieslik 2005a).

Our dependent variable is a count measuring the number of firms (entities of the national economy recorded in the firms' REGON register) operating in a given region. These firms are enrolled in the Polish commercial register, the Krajowy Rejestr Sądowy.

To explain the location decision behavior of firms across different regions of the Polish economy, we follow the past literature (as described in the previous chapter) in applying a set of standard determinants at the regional level: GDP, industry agglomeration economies, services agglomeration economies, urbanization economies, the wage level, the unemployment rate, human capital, the land area and the quality of the infrastructure. We additionally control for R&D agglomeration economies and patents applications to capture innovative activities. Moreover, we control for accession to the European Union and for border effects. We will further measure the contingent effects of European Union membership by adding interaction terms.

The value of the region-level GDP is on the one hand intended to capture the economic size of market demand. A higher level of regional GDP, leading to higher consumer expenditures in that region, can be expected to provide an incentive for firms to choose that location.[5] GDP might also indicate the size of the supplier side and indicate a large pool of intermediate goods suppliers.

As noted in the previous section, different agglomeration factors could play a relevant role in explaining firms' location choices. First, we measure agglomeration economies separately for industrial and service sectors. They are expressed as the respective shares of employment in industry and services over the total employment in the region. The stronger the presence of industrial and/or service sectors in a region, the better represented potential suppliers are and consequently the higher the expected benefits from locating in that region.

Moreover, the past literature recognized that urbanization economies might play an important role in determining location choices, but the direction of influence is still a matter of discussion. On the one hand, urbanization might yield benefits via labor market pooling, improved infrastructure and better information networks. On the other hand, dense agglomeration might also generate congestion costs, deriving, for example, from pollution, traffic jams or social distress. Such negative factors might consequently discourage locating in highly urbanized regions. Urbanization effects are measured in a standard way, as a percentage share of urban population over total population in a region.

As a novelty to the agglomeration literature for the new member states, we implement new agglomeration factors, referring more precisely to local innovative efforts and locally available knowledge. In

particular, we introduce a variable measuring agglomeration economies stemming from the R&D sector. This variable is constructed as the share of the R&D sector's employment over the total region's employment. The R&D sector is considered a measure of the region's overall effort to enhance the local degree of innovativeness (Mudambi and Swift 2012). Given, however, the technological and commercial uncertainty to actually transform such innovative input into an innovative and marketable output, R&D activities do not necessarily reflect the readily available knowledge base. This notwithstanding, firms observing a relatively high R&D intensity in a region should be encouraged to locate. Instead, to approximate the actual innovative output of a region, we include the number of patents: The higher the number of patents in a region, the more attractive for the firms it is to locate and to take advantage from potential knowledge spillovers.[6]

Further variables are designed to capture labor market effects. In particular, a higher wage level induces higher labor costs for potential employers and will consequently reduce their incentive to open a business in a region. At the same time, higher wages might act as an indicator of the quality of the labor force. Thus, the net effect of the wage level on the location decision is theoretically unclear. The measure we take is average monthly gross wages and salaries.

Another labor market factor is the regional unemployment rate. High unemployment might indicate a region's economic decline and as such will be detrimental to firms' location decisions. Our measure here is the average annual unemployment rate for the overall active population.

Finally, the share of students per population serves as a measure of a region's educational level. A higher share of students indicates a region's ability to increase the potential workforce's qualification and degree of knowledge spillovers. A firm deciding where to locate will judge positively on this regional attribute and will be more likely to choose that region. With our indicator, we follow the literature, measuring the stock of human capital in terms of labor force that completed secondary and/or tertiary levels of education (Alama-Sabater et al. 2011; Coughlin and Segev 2000).

The quality of the infrastructure might also positively influence location choices. More precisely, improvement of infrastructure increases the economic potential of regions and enhances agglomeration forces (Cieślik and Rokicki 2013). Moreover, a high quality of infrastructure reduces transportation time and distance, helping firms reach both suppliers and consumers. According to the literature, the quality of infrastructure is measured by the density of the road network, railway lines network, share of telephone lines per population and a region's area. A positive influence for the road network, railway lines and telephone lines on the firms' decision to locate can be expected, but the influence of a region's area is not clear. On the one hand, a larger area

will offer more space to build production plants. On the other hand, a larger area implies that the transportation network density will be smaller.

Furthermore, to capture the influence of unobservable factors referring to regional characteristics, we introduce regional dummy variables (for the regions in the North-West, North, South-West, South and East, where the central region is the reference category), as well as time dummies to account for time-specific effects. Alternatively to the regional dummies, we introduce a measure capturing the distance to the German border, measured as the distance of a Polish region's capital city on main car routes to a German border city: Penkun, Goerlitz or Frankfurt an der Oder. We extracted this information from Google maps. Finally, we took a dummy variable capturing the time of Poland's accession to the EU, counting a zero until 2003 and a one thereafter.

Methodology

Given that the dependent variable in our regression framework is a count, taking exclusively nonnegative integer values, we will have to consider count-data regression methods. In econometric theory, count data are mostly modeled by a Poisson distribution. Other possible methods are the binomial or negative binomial distribution. The benefit of modeling a binomial distribution is the capability to deal with an upper bound of counts, whereas by modeling a negative binomial distribution conditional moments restrictions imposed by the Poisson distribution can be eased.

Under the assumption of a Poisson distribution for a given count variable *y* and a given set of explanatory variables *X,* the density can be written as (Wooldridge 2002):

$$f(y|X) = \exp[-\mu(X)]\,[\mu(X)]^{y}/y! \qquad (1)$$

where $\mu(X) \equiv E(y|X)$ denotes the conditional mean. A conditional maximum likelihood estimator can be derived and the estimator will be efficient. A strong restriction, however, is imposed on the conditional moments: The conditional variance and mean are assumed to be equal.

In our context, the number of firms y_i is drawn from a Poisson distribution with parameter $\mu\,(x_{\,i})$ and is dependent on a set of regional variables x_i, contained in the vector X_i,such that the probability to observe a count of firms is:

$$P(y_{\,i}|X_{\,i}) = \exp[-\mu(X_{\,i})]\,[\mu(X_{\,i})]^{y_{\,i}}/y_{\,i} \qquad (2)$$

The most common functional form taken for the mean is the exponential function, such that $\mu(X_{\,i}) = \exp(X_{\,i}\beta)$. Vector β is the column vector of coefficients measuring the impact of regional explanatory variables and has to be estimated.

To circumvent the problematic moment restrictions for Poisson models, econometric theory offers the Negbin II model of Cameron and Trivedi (1986), which is a negative binomial regression model. In this model, an additional term c_i capturing unobserved heterogeneity is considered. It is assumed that c_i is independent of x_i and has a Gamma distribution with unit mean and variance η^2. The conditional mean is the same as in the Poisson model, but the variance is different (bigger than the mean) and can be written as:

$$Var(y_i|X_i) = E(y_i|X_i) + \eta^2 (E(y_i|X_i))^2 \quad (3)$$

The panel structure of our data creates a need to control for unobserved heterogeneity. For that reason, we opt for negative binomial estimation as our preferred estimation technique.

Descriptive evidence

Taking a look at the spatial distribution of firms in the Polish economy in 2010 as shown in Table 1, reveals that across regions, the largest number of firms is given for the regions Mazowieckie –the central area around the Polish capital city Warsaw— Śląskie, Wielkopolskie, Malopolskie and Dolnośląskie. Clearly, the Southern, Western and Central parts around the capital city bear the highest firm activity in the Polish economy. These are also regions with economically important urban centers like Katowice in Śląskie, Krakow in Malopolskie, Poznań in Wielkopolskie and Wrocław in Dolnośląskie. At the same time, Mazowieckie remains the leader both in terms of the number of firms operating in the region and of the relative volume of investment.[7]

The greatest dynamics in terms of growth of the number of firms are found for the regions Mazowieckie, Malopolskie, Wielkopolskie and Pomorskie (with the three important harbor cities Gdańsk, Gdynia and Sopot). The data reveal less growth of the number of firms in the Eastern parts of the country between 2003 and 2010. These regions continue to experience the lowest GDP per capita relative to the Polish average. In 2010, GDP per capita was still between 67% of the Polish average in Podkarpackie and Lubelskie and 73% in Podslaskie and Warmińsko-Mazurkie, compared to 162% in Mazowieckie (GUS 2013).

--Insert Table 1 about here--

Table 2 demonstrates that Poland scored relatively low in terms of R&D expenditures per capita and innovative activities as measured by patent applications to the European Patent Office (EPO) in the beginning of the 21st century. The country ranked at position 4 or 3, respectively, behind the recent new EU member states Romania and Bulgaria. Compared to the old EU member states like the Scandinavian countries Sweden,

Denmark, and Finland, Poland had only about 1-3% of their value in R&D expenditures or patent applications, respectively. However, as the figures show, until 2010, Poland made considerable progress and was only outweighed by the impressive progress made by Estonia, Latvia and Czech Republic and in terms of R&D expenditures by Malta. However, Poland still has a long way to go before it will meet its R&D targets according to the European 2020 strategy. One target is to increase R&D expenditures to 1.7% of its national GDP. To reach the goals, in 2010, Poland set up a reform of six legislative acts titled "Building upon Knowledge: Science Reform for Poland's Development" to provide financial funding for research institutions. Moreover, under the "National Reform Programme" set up in 2011, working conditions in the R&D sector will be promoted, the higher education sector will be modernized and quality control of research and education will be implemented.

--Insert Table 2 about here--

ECONOMETRIC ANALYSIS

The empirical analysis examines the location choice of firms in the Polish economy from 2003 to 2010. We investigate the determinants of the location of firms using a negative binomial regression analysis.[8] More specifically, the expected count of firms will be investigated with regard to regional explanatory characteristics. In the regression analyses, we will use lagged explanatory variables to avoid simultaneity problems. We differentiate the effects between foreign firms, domestic firms and entrepreneurial firms in the Polish economy.

Foreign Firms

The prior literature has mainly focused on the activities and location of multinational enterprises and FDI. Cieślik (2005a), for example, investigated the location decisions of firms with foreign capital participation in the Polish economy for the period from 1993 to 1998 using regional NUTS II data.[9] In our analysis, we used the same data source, but we capture a more recent time period from 2003 to 2010. This period is important because it captures the very first years of Poland being a member of the EU.[10] In a similar vein as in Cieślik's study, our analysis was restricted by data availability, so we cannot control for a degree of foreign capital participation that would characterize the status of FDI, which is commonly given by a share of more than 10%. Instead, we have to deal with all firms that have at least some degree of foreign capital participation.

--Insert Table 3 about here--

Our results show that all sectors' agglomeration economies are important for the location decision of multinational enterprises, supporting evidence found by Cieślik (2005a). The results suggest that the presence of a broad local basis of services is decisive when it comes to defining business strategies and making location choices. Moreover, agglomeration economies emerging from the R&D sector play a significant role.

In contrast to Cieslik, we can find significant effects stemming from market demand and the quality of human capital. The literature shows that the choice of the variable for representing the degree of human capital does, in fact, matter (Arauzo-Carod 2013). Our measure for human capital is different from Cieślik's. We further establish a positive effect from locating closer to the German-Polish border. Moreover, the results point to urbanization diseconomies and a negative impact resulting from unemployment. Most importantly, our results imply that the intensity of local innovative activity forms a central force for the location decision of multinational enterprises: the coefficient on patents is significant and positive. Interestingly, the coefficients for the interaction terms indicate that due to accession to the European Union, a stronger impact for locating closer to the German-Polish border occurred, as well as a positive impact due to innovative activity.

Domestic firms

The results for domestic firms demonstrate that industrial and R&D agglomeration economies significantly influence the firms' choice to locate in a given region. However, no clear or significant effect stems from services agglomeration economies. A positive impact results from market demand and a negative one from wages. The quality of human capital as measured by the share of students per region bears a significant positive impact. We further find that the quality of the road network positively and significantly influences the number of firms in a region. Road transportation is the only infrastructure that positively determines the location of firms; the effect for railways and telephone lines in the majority of cases remained insignificant.

--Insert Table 4 about here--

The validity of the negative binomial regression model can be assessed by looking at the value of alpha that results from a regression testing for equidispersion (Cameron and Trivedi 2009). Equidispersion implies that the conditional variance is not larger than the conditional mean and there would be no reason to estimate the negative binomial model. In other words, if alpha equals zero, the Poisson regression model is preferred over the

negative binomial regression model. Our results imply that the negative binomial regression model is the adequate choice, since in every specification we reject the null hypothesis of equidispersion.

New Firms

As we illustrated in the theory section, there are reasons to expect that factors determining firms' location choice work differently for newly established firms. Generally, the estimation results with the dependent variable given by the count of newly registered firms confirm the previous findings, namely, that industry and R&D agglomeration economies as well as the share of students per population and road networks are important explanatory factors. There are some weaker effects due to services agglomeration economies and market demand. This suggests that a higher market demand is less important for the regional location choice of new firms that are just entering the market. It can be assumed that it is crucial for already established firms to meet a sufficient market size for their survival. For entering firms, it can well be that the market is still under formation, so that its initial size is not significantly determining their decision to enter. The effect due to EU membership is positively significant. When estimating the regressions for newly registered firms, we are grasping the effect of firms deciding to locate, so opportunities to exploit the enlarged common market could have motivated more new establishments in the Polish regions. The EU membership has thus increased the attractiveness of the Polish regions for new investors. Finally, accession to the European Union bore a positive impact for the effect of innovative activity as measured by patents applications for the location decision of new firms. Similarly as before, from the results of the alpha-test, we conclude that negative binomial estimators are preferred over Poisson estimators.

--Insert Table 5 about here--

Tests for Spatial Autocorrelation

Another concern in location studies could arise from estimation biases due to spatial autocorrelation. The inherent problem is that the spatial distribution of a variable might be not random but instead will show a systematic pattern that will violate the independence assumption of many estimation procedures. OLS estimation, for example, is dependent on the observations being chosen randomly and independently from each other. To test for spatial autocorrelation we compute Moran's I (see Moran 1950), which is a well-known indicator in the literature. This procedure requires setting up a spatial weighting matrix that captures the distance between two observation units. For our analysis, we took the inverse of the distances of the 16 Polish

voivodeships' capital cities to each other by km distance on main car routes. As such, observation units that are closer to each other attain a higher value. As Table 6 shows, the null hypothesis of zero spatial autocorrelation cannot be rejected for the count of all firms in the Polish economy. Since the z-scores are very low and Moran's I value is close to zero, we can safely conclude that spatial autocorrelation is not a problem in our study.

--Insert Table 6 about here--

CONCLUSIONS

In this paper we focused on the location decisions of firms in Poland. In our analysis, we discovered rich firm dynamics in the Polish economy and figured out the differential drivers for the location decision of domestic firms, foreign firms and entrepreneurs in the economy. The former literature, however, focused predominantly on the location decision of FDI. Moreover, the prior literature lacks an investigation of the impact of agglomeration economies stemming from the generation and diffusion of knowledge: the R&D sector activities and the innovative activities.

The analysis in this paper has shown that the location choice of firms in Poland is decisively influenced by agglomeration economies stemming from the R&D sector, as well as by the quality of human capital. Given that R&D expenditures and innovative activity of Poland are low compared to the other European Union member countries, for the future, a stronger focus on supporting these activities will be important to attract firms' location. Policies aimed at sustaining proper training and skill upgrading of the local labor force as well as at stimulating a dynamic R&D environment and innovative activities will offer favorable conditions for the movement of goods and production inputs and create an incentive for the location decision of firm activity.

The study revealed different effects that foster the location decision of either domestic, foreign or entrepreneurial firm activity. Market demand is particularly important for the domestic and foreign firms' location decision but does not play a role for new firms which focus on the quality of the infrastructure and human capital. Consequently, in addition to improving education and training, improving infrastructure should remain an important policy priority to attract new firms to a region. For newly registered firms, positive effects due to accession to the European Union could be detected. Regarding multinational enterprises, our analysis demonstrated that due to accession to the European Union, firms tend to locate closer to the German-Polish border. Moreover, the impact of innovative activity became stronger after European Union accession.

Regarding the European structural funds, in the period 2007 to 2013, Poland received the largest share of financial means. All firms located in Poland were able to apply for EU funding. These finances were to be

employed for the qualification of human capital, innovative investments and cooperation between research institutes and enterprises, among other uses. Given that Poland's accession to the European Union bears positive effects for the establishment of new firm activity, the transition from the planned economy to an open market economy and finally to becoming a member in the European market can be seen as a dynamic process that brought about competitive advantages for improving human capital and innovation. However, the comparatively low levels of R&D expenditures and innovative activities need to be further improved to provide favorable conditions for the geographical location decision of firms in the Polish economy and to reach the country's R&D targets with regard to the European 2020 strategy goals.

ACKNOWLEDGEMENTS

We are grateful for comments from Maite Alguacil, Carsten Eckel, Barbara Liberda, Admasu Shiferaw, the participants of the Third Euroacademia International Conference Re-Inventing Eastern Europe, participants of the Goettingen Workshop International Economic Relations and to participants of the First World Congress in Comparative Economics.

REFERENCES

Alamá-Sabater, L., A. Artal-Tur, and J. M. Navarro-Azorín. 2011 "Industrial location, spatial discrete choice models and the need to account for neighbourhood effects" *Annals of Regional Science* 47, 393–418.

Arauzo-Carod, J. M. 2013. "Location determinants of new firms: Does skill level of human capital really matter?" *Growth and Change* 44, no. 1: 118–48.

Arauzo-Carod, J. M., D. Liviano-Solis, and M. C. Manjon-Antolin. 2010. "Empirical studies in industrial location: An assessment of their methods and results" *Journal of Regional Science* 50, no. 3: 685–711.

Arauzo-Carod, J. M., and E. Viladecans-Marsal. 2009. "Industrial location at the intra-metropolitan level: The role of agglomeration economies" *Regional Studies* 43, no. 4: 545–58.

Arauzo-Carod, J. M., and M. C. Manjon-Antolin. 2004. "Firm size and geographical aggregation: An empirical appraisal in industrial location" *Small Business Economics* 22, no. 3-4: 299–312.

Bade, F. J., and E. A. Nerlinger. 2000. "The spatial distribution of new technology-based firms: Empirical results for West-Germany" *Papers in Regional Science* 79, no. 2: 155–76.

Bartik, T. 1985. "Business location decisions in the United States: estimates of the effects of unionization, taxes, and other characteristics of states" *Journal of Business and Economic Statistics* 3, 14–22.

Broadman, H. G., and X. Sun. 1997. “The distribution of foreign direct investment in China” *World Economy* 20, no. 3: 339–61.

Cameron, A. C., and P. K. Trivedi. Microeconometrics Using Stata. College Station: Stata Press, 2009.

Cameron, A. C., and P. K. Trivedi. 1986. “Econometric models based on count data: Comparisons and application of some estimators and tests” *Journal of Applied Econometrics* 1, no. 1: 29–54.

Carlin, W., J. van Reenen, and T. Wolfe. 1995. “Enterprise restructuring in early transition: The case study evidence from central and eastern Europe” *Economics of Transition* 3, no. 4: 427–58.

Carlton, D. W. 1983. “The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables” *Review of Economics and Statistics* 65, no. 3: 440–49.

Chidlow, A., L. Salciuviene, and S. Young. 2009. “Regional determinants of inward FDI distribution in Poland” *International Business Review* 18, no. 2: 119–33.

Cieślik, A. 2005 a. “Regional characteristics and the location of foreign firms within Poland” *Applied Economics* 37, no. 8: 863–74.

Cieślik, A. 2005 b. “Location of foreign firms and national border effects: The case of Poland” *Tijdschrift voor Economische en Sociale Geografie* 96, 287–97.

Cieślik, A., and B. Rokicki. 2013. “Rola sieci transportowej w rozwoju polskich regionow: zastosowanie modelu potencjału ekonomicznego” *Acta Universitatis Nicolai Copernici* 44, 113–27.

Coughlin, C. C., and E. Segev. 2000. “Location determinants of new foreign-owned manufacturing plants” *Journal of Regional Science* 40, 323–51.

Crozet, M., T. Mayer, and J. L. Mucchielli. 2004. “How do firms agglomerate? A study of FDI in France” *Regional Science and Urban Economics* 34, 27–54.

Disdier, A., and T. Mayer. 2004. “How different is Eastern Europe? Structure and determinant of location choices by French firms in Eastern and Western Europe” *Journal of Comparative Economics* 32, no. 2: 280-96.

Fidrmuc, J. P. 2007. “Channels of restructuring in privatized Czech companies” *Economics of Transition* 15, 309–39.

Figueiredo, O., P. Guimarães, and D. P. Woodward. 2002. “Home-field advantage: location decisions of Portuguese entrepreneurs” *Journal of Urban Economics* 52, 341–61.

Gauselmann, A., and P. Marek. 2012. “Regional determinants of MNE’s location choice in transition economies” *Empirica* 39, 487–511.

Glaeser, E. L. 1999. “Learning in cities” *Journal of Urban Economics* 46, 254–77.

Glickman, N. J., and D. P. Woodward. 1988. "The location of foreign direct investment in the United States: patterns and determinants" *International Regional Science Review* 11, 137–54.

Goldstein, G., and T. Gronberg. 1984. "Economies of scope and economies of agglomeration" *Journal of Urban Economics* 16, 91–104.

Guimarães, P., O. Figueiredo, and D. Woodward. 2000. "Agglomeration and the location of foreign direct investment in Portugal" *Journal or Urban Economics* 47, 115–35.

GUS (Glowny Urzad Statystyczny). Wpływ potencjału demograficznego i gospodarczego miast wojewódzkich na kondycję województw. Warsaw Statistical Office, 2013.

Head, K., and J. Ries. 1996. "Inter-city competition for foreign investment: static and dynamic effects of China's incentive areas" *Journal of Urban Economics* 40, 38–60.

Head, K., J. Ries, and D. L. Swenson. 1995. "Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States" *Journal of International Economics* 38, 223–47.

Helsley, R. W., and W. C. Strange. 1990. "Matching and agglomeration economies in a system of cities" *Regional Science and Urban Economics* 20, 189–212.

Hilber, C. A. L., and I. Voicu. 2010. "Agglomeration economies and the location of foreign direct investment: Empirical evidence from Romania" *Regional Studies* 44, 355–71.

Holl, A. 2004. "Start-ups and relocations: Manufacturing plant location in Portugal" *Papers in Regional Science* 83, 649–68.

Krugman, P. 1991. "Increasing returns and economic geography" *Journal of Political Economy* 99, 483–99.

Krugman, P. 1996. "Urban concentration: the role of increasing returns and transport costs" *International Regional Science Review* 19, 5–30.

Luger, M. I., and D. Shetty. 1985. "Determinants of foreign plant start-ups in the United States: lessons for policy makers in the Southeast" *Vanderbilt Journal of Transnational Law* 18, 223–45.

Manjon-Antolin, M. C., and J. M. Arauzo-Carod. 2011. "Locations and relocations: Determinants, modeling, and interrelations" *Annals of Regional Science* 47, no. 1: 131–46.

Marshall, A. Principles of Economics. London: Macmillan, 1890.

Moran, P. 1950. "Notes on Continuous Stochastic Phenomena" *Biometrika* 37, 17–23.

Mudambi, R., and T. Swift. 2012. "Multinational enterprises and the geographical clustering of innovation" *Industry and Innovation* 19, 1–21.

Pusterla, F., and L. Resmini. 2007. "Where do foreign firms locate in transition countries? An empirical investigation" *Annals of Regional Science* 41, no. 4: 835-56.

Quigley, J. 1998. “Urban diversity and economic growth” *Journal of Economic Perspectives* 12, 127–38.

Robinson, I. 2004. “Enterprise restructuring in transition: Case study evidence from Poland” *Journal of East-West Business* 10, 45–64.

Rosenthal, S. S., and W. C. Strange. 2001. “The determinants of agglomeration” *Journal of Urban Economics* 50, 191–229.

Smith, D. F., and R. Florida. 1994 “Agglomeration and industrial location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries” *Journal of Urban Economics* 36, 23–41.

Vechiu, N., and F. Makhlouf. 2014. “Economic integration and specialization in production in the EU27: does FDI influence countries’ specialization?” *Empirical Economics* 46, 543-72.

Woodward, D. P. 1992. “Locational determinants of Japanese manufacturing start-ups in the United States” *Southern Economic Journal* 58: 690–708.

Wooldridge, J. Econometric Analysis of Cross Section and Panel Data. MA: MIT Press, 2002.

TABLE 1
The spatial distribution of firm activity in the Polish economy

Year	NUTS II region	Location	Firms with foreign capital participation	Domestic firms	New firms	All firms	Growth of new firms from 2003 to 2010
2010	Dolnośląskie	South-West	2275	328832	33257	331107	55,06
	Kujawsko-pomorskie	North	567	185379	19518	185946	49,80
	Lubelskie	East	349	163701	17647	164049	46,43
	Lubuskie	North-West	778	105330	11767	106107	55,53
	Łódzkie	Central	941	229912	25170	230853	46,84
	Małopolskie	South	1468	329866	36370	331334	77,47
	Mazowieckie	Central	8585	672447	66365	681032	67,82
	Opolskie	South-West	468	98611	8545	99079	52,72
	Podkarpackie	East	340	152233	15960	152574	52,20
	Podlaskie	East	144	91679	10229	91823	51,58
	Pomorskie	North	1252	258837	27533	260089	61,26
	Śląskie	South	2086	449357	46349	451443	58,55
	Świętokrzyskie	East	180	108580	10517	108760	50,79
	Warmińsko-mazurskie	North	305	118759	13666	119065	49,43
	Wielkopolskie	North-West	2068	373282	38259	375351	67,45
	Zachodniopomorskie	North-West	1293	219174	20857	220466	46,40
2003	Dolnośląskie	South-West	1681	304091	21448	305772	-
	Kujawsko-pomorskie	North	414	190683	13029	191096	-
	Lubelskie	East	307	154609	12051	154916	-
	Lubuskie	North-West	595	96753	7566	97348	-
	Łódzkie	Central	649	240880	17141	241530	-
	Małopolskie	South	813	287073	20494	287886	-
	Mazowieckie	Central	4930	570785	39545	575716	-
	Opolskie	South-West	422	86987	5595	87409	-
	Podkarpackie	East	252	142361	10486	142613	-
	Podlaskie	East	84	96807	6749	96891	-
	Pomorskie	North	897	225437	17074	226334	-
	Śląskie	South	1603	422274	29233	423877	-
	Świętokrzyskie	East	168	103547	6975	103715	-
	Warmińsko-mazurskie	North	286	110167	9145	110453	-
	Wielkopolskie	North-West	1310	334347	22848	335657	-
	Zachodniopomorskie	North-West	967	199679	14247	200645	-

Source: Statistical Regional Yearbooks of the Polish Central Statistical Office, authors' computations.

Note: The Table displays the spatial distribution of firms with foreign capital participation, domestic firms, entrepreneurial firms as well as all firms in the economy for the years 2003 and 2010. Moreover, the growth of new firms from 2003 until 2010 is shown.

TABLE 2
R&D expenditures and innovative activity for the EU27

Ranking of countries according to their degree of patent applications and R&D expenditures

Patent applications to the EPO; per million inhabitants			Total intramural R&D expenditures; in Euros per inhabitants		
Country	2003	2010	Country	2003	2010
Romania	0,747	1,602	Romania	9,4	28,2
Bulgaria	2,712	2,287	Bulgaria	11,4	29
Poland	2,86	9,375	Latvia	16,4	51,2
Latvia	2,983	7,442	Poland	27,1	68,6
Lithuania	4,91	5,061	Malta	28,8	101,9
Slovakia	5,84	8,626	Slovakia	31,5	77,2
Portugal	6,388	8,898	Lithuania	32,2	69,9
Cyprus	7,048	9,363	Estonia	48,6	174,6
Estonia	7,563	28,073	Cyprus	57,4	105,2
Greece	7,696	5,839	Hungary	68,3	112,4
Malta	8,18	8,454	Greece	88,9	120,9
Czech Republic	10,87	18,235	Portugal	97,6	260,8
Hungary	13,021	19,286	Czech Republic	99,3	200,3
Spain	23,015	32,098	Slovenia	164,2	364,4
Slovenia	36,811	50,772	Spain	196,4	313,8
Ireland	56,902	68,787	Italy	258,5	331,6
Italy	76,903	75,28	Ireland	412,9	586,8
United Kingdom	95,056	83,746	United Kingdom	483,3	491,6
France	128,574	130,387	Belgium	500	690,7
Belgium	131,402	138,506	France	558,8	672,3
Austria	171,403	209,396	Netherlands	565	657,1
Luxembourg	194,446	150,239	Austria	622,4	965,9
Denmark	206,189	228,793	Germany	660,8	855,1
Netherlands	215,89	182,197	Denmark	901,8	1.281,6
Sweden	230,095	296,158	Luxembourg	949,8	1.178,3
Finland	248,659	256,636	Finland	961,3	1.302,7
Germany	268,794	283,54	Sweden	1.186,3	1.270,8

Source: Eurostat.

Note: EPO=European Patent Office. The Table shows the R&D expenditures and the innovative activity measured in terms of patent applications to the EPO for the EU27 countries for the years 2003 and 2010.

TABLE 3
Regression results for firms with foreign capital participation

Negative binomial estimation				
VARIABLES	(1)	(2)	(3)	(4)
Constant	2.380	4.345***	1.294689*	4.329***
	(1.500)	(1.247)	(0.7557)	(1.338)
GDP	7.70e-06***	9.29e-06***	7.31e-06***	9.26e-06***
	(2.34e-06)	(2.31e-06)	(2.00e-06)	(2.40e-06)
Industry agglom. econ.	0.0358*	0.0304*	0.0641***	0.0299
	(0.0194)	(0.0178)	(0.0151)	(0.0184)
Services agglom. econ.	0.0711***	0.0381***	0.0599***	0.0369**
	(0.0144)	(0.0138)	(0.0161)	(0.0149)
Urbanization econ.	-0.0252**	-0.0198**	-0.0255***	-0.0186*
	(0.0124)	(0.00996)	(0.0087)	(0.0105)
Wages	-0.00101**	-0.000467*	-0.00036	-0.000455
	(0.000450)	(0.000269)	(0.00028)	(0.000291)
Unemployment	-0.0456***	-0.0256**	-0.02298**	-0.0204**
	(0.0158)	(0.0102)	(0.0109)	(0.00958)
Students	0.352***	0.227***	0.2767***	0.201***
	(0.102)	(0.0597)	(0.0731)	(0.0586)
Telephone lines	0.0128	0.00701	0.01896	0.00479
	(0.0129)	(0.0176)	(0.01947)	(0.0178)
Road network	0.000300	-0.00581	-0.0046	-0.00513
	(0.00692)	(0.00374)	(0.0034)	(0.00375)
Railway lines	0.139*	0.0847*	0.0694	0.0784
	(0.0743)	(0.0480)	(0.0433)	(0.0495)
Area		1.07e-05	0.00002*	1.17e-05
		(1.18e-05)	(0.000012)	(1.21e-05)
R&D agglom. econ.	1.343***	1.012***	0.66006***	1.124***
	(0.334)	(0.304)	(0.18957)	(0.304)
Patents	0.00784	0.0167***	0.021007***	
	(0.00802)	(0.00483)	(0.0077)	
EU membership		-0.0816*		
		(0.0455)		
GE-PL border		-0.00158***		-0.00161***
		(0.000398)		(0.000423)
Interact EU-border			-0.00063***	
			(0.0002)	
Interact. EU-patents				0.0123**
				(0.00501)
Time dummies	yes	no	no	no
Regional dummies	yes	no	no	no
Log likelihood	-767.63206	-768.91505	-792.0164	-771.5979
alpha	0.02701	0.02762	0.04099	0.029068
s.e.	0.0058	0.0063	0.0132	0.0065
Observations	123	123	123	123

Source: Authors' computations based on the Statistical Regional Yearbooks of the Polish Central Statistical Office and Eurostat.

Note: The results refer to negative binomial estimates on the pooled data for the sample of firms with foreign capital participation. Cluster-robust standard errors are displayed in parentheses. The Table reports the log likelihood, the value of alpha resulting from a regression testing for equidispersion which gives a value of zero if the Poisson regression model is preferred over the negative binomial regression model. * denotes significance at a 10% level, ** denotes significance at a 5% level, *** denotes significance at a 1% level.

TABLE 4
Regression results for domestic firms

Negative binomial estimation				
VARIABLES	(1)	(2)	(3)	(4)
Constant	10.08***	9.560***	9.647***	9.548***
	(0.928)	(0.971)	(0.465)	(0.990)
GDP	8.14e-06***	3.61e-06**	3.65e-06***	3.55e-06**
	(1.58e-06)	(1.50e-06)	(1.40e-06)	(1.52e-06)
Industry agglom. econ.	0.0320***	0.0281*	0.0259***	0.0272*
	(0.0118)	(0.0146)	(0.00596)	(0.0149)
Services agglom. econ.	0.0336***	0.00825	0.00713	0.00733
	(0.0108)	(0.0133)	(0.00918)	(0.0137)
Urbanization econ.	0.00153	0.00611	0.00631	0.00696
	(0.00836)	(0.00745)	(0.00774)	(0.00732)
Wages	-0.00129***	-0.000271*	-0.000258	-0.000271*
	(0.000261)	(0.000150)	(0.000164)	(0.000157)
Unemployment	-0.0371***	0.00148	0.00229	0.00413
	(0.0123)	(0.00550)	(0.00465)	(0.00511)
Students	0.291***	0.0895***	0.0832**	0.0780**
	(0.0503)	(0.0340)	(0.0360)	(0.0339)
Telephone lines	-0.00531	-0.0121	-0.0126	-0.0133*
	(0.0161)	(0.00796)	(0.00994)	(0.00743)
Road network	0.00584	0.00652**	0.00669**	0.00675**
	(0.00526)	(0.00307)	(0.00311)	(0.00304)
Railway lines	0.0346	-0.00656	-0.00696	-0.00764
	(0.0440)	(0.0280)	(0.0281)	(0.0282)
Area		3.23e-05***	3.22e-05***	3.31e-05***
		(1.09e-05)	(9.54e-06)	(1.07e-05)
R&D agglom. econ.	0.894***	0.528**	0.541**	0.565**
	(0.224)	(0.215)	(0.258)	(0.221)
Patents	-0.0124	-0.00203	-0.00266	
	(0.00783)	(0.00679)	(0.00788)	
EU membership		-0.0540*		
		(0.0307)		
GE-PL border		-3.14e-06		-4.10e-06
		(0.000461)		(0.000473)
Interact. EU-border			-0.000110	
			(0.000140)	
Interact EU-patents				-0.00144
				(0.00548)
Time dummies	yes	no	no	no
Regional dummies	yes	no	no	no
Log likelihood	-1418.1553	-1432.4402	-1432.2718	-1433.4273
alpha	0.01528	0.01925	0.0192	0.01956
s.e.	0.00277	0.0059	0.0058	0.00615
Observations	123	123	123	123

Source: Authors' computations based on the Statistical Regional Yearbooks of the Polish Central Statistical Office and Eurostat.

Note: The results refer to negative binomial estimates on the pooled data for the sample of domestic firms. Cluster-robust standard errors are displayed in parentheses. The Table reports the log likelihood, the value of alpha resulting from a regression testing for equidispersion which gives a value of zero if the Poisson regression model is preferred over the negative binomial regression model. * denotes significance at a 10% level, ** denotes significance at a 5% level, *** denotes significance at a 1% level.

TABLE 5
Regression results for new firms

Negative binomial estimation				
VARIABLES	(1)	(2)	(3)	(4)
Constant	7.627***	6.241***	6.763***	6.167***
	(0.907)	(1.073)	(0.538)	(1.053)
GDP	8.93e-06***	1.59e-06	1.94e-06*	1.18e-06
	(1.62e-06)	(1.16e-06)	(1.14e-06)	(1.19e-06)
Industry agglom. econ.	0.0384***	0.0354**	0.0304***	0.0359**
	(0.0128)	(0.0146)	(0.00648)	(0.0142)
Services agglom. econ.	0.0356***	0.0167	0.0140*	0.0161
	(0.00999)	(0.0134)	(0.00787)	(0.0133)
Urbanization econ.	0.00195	0.0128	0.0129	0.0128
	(0.00873)	(0.00882)	(0.00952)	(0.00875)
Wages	-0.00158***	-0.000247	-0.000267	-0.000214
	(0.000297)	(0.000161)	(0.000171)	(0.000152)
Unemployment	-0.0345***	0.000490	-0.00215	0.000871
	(0.0108)	(0.00664)	(0.00593)	(0.00635)
Students	0.346***	0.0831**	0.0839*	0.0746*
	(0.0586)	(0.0409)	(0.0488)	(0.0413)
Telephone lines	-0.00666	-0.0336***	-0.0343***	-0.0320***
	(0.0145)	(0.00940)	(0.0111)	(0.00906)
Road network	0.00125	0.00717**	0.00670*	0.00735**
	(0.00482)	(0.00362)	(0.00375)	(0.00363)
Railway lines	0.0449	-0.0167	-0.0120	-0.0160
	(0.0466)	(0.0348)	(0.0339)	(0.0345)
Area		4.28e-05***	4.08e-05***	4.44e-05***
		(1.17e-05)	(1.05e-05)	(1.13e-05)
R&D agglom. econ.	1.012***	0.625**	0.653**	0.644**
	(0.231)	(0.258)	(0.294)	(0.250)
Patents	-0.00622	0.0108	0.00984	
	(0.00664)	(0.00675)	(0.00727)	
EU membership		0.0613**		
		(0.0286)		
GE-PL border		0.000273		0.000287
		(0.000465)		(0.000457)
Interact. EU-border			0.000133	
			(0.000144)	
Interact EU-patents				0.0166***
				(0.00619)
Time dummies	yes	no	no	no
Regional dummies	yes	no	no	no
Log likelihood	-1111.2211	-1136.6904	-1138.2087	-1135.9054
alpha	0.0154	0.0233	0.0239	0.02305
s.e.	0.00229	0.0068	0.00677	0.0065
Observations	123	123	123	123

Source: Authors' computations based on the Statistical Regional Yearbooks of the Polish Central Statistical Office and Eurostat.

Note: The results refer to negative binomial estimates on the pooled data for the sample of new firms. Cluster-robust standard errors are displayed in parentheses. The Table reports the log likelihood, the value of alpha resulting from a regression testing for equidispersion which gives a value of zero if the Poisson regression model is preferred over the negative binomial regression model, and the respective robust standard errors. * denotes significance at a 10% level, ** denotes significance at a 5% level, *** denotes significance at a 1% level.

TABLE 6
Tests for spatial autocorrelation

	year	Moran's I	z-score	p-value
All firms	2003	-0.062	0.109	0.457
	2004	-0.063	0.072	0.471
	2005	-0.06	0.144	0.443
	2006	-0.059	0.165	0.435
	2007	-0.058	0.183	0.427
	2008	-0.056	0.234	0.407
	2009	-0.06	0.154	0.439
	2010	-0.062	0.102	0.459

Source: Authors' computations based on the Statistical Regional Yearbooks of the Polish Central Statistical Office.

Note: This Table displays results from the computation of Moran's I test for spatial autocorrelation. For the spatial weighting matrix we took the inverse of the distances of the 16 Polish voivodeships' capital cities to each other, given by main car routes. This information was collected from Google maps. All the Moran's I values, z-scores and the corresponding p-values are in favor of indicating no spatial autocorrelation.

APPENDIX

TABLE A1
List of variables

Variable	Description (detailed way of extraction from the Polish Central Statistical Office and other sources)
Firms, all	Entities of the economy entered, entities of the national economy-indicators, entities-indicators, entities entered in the Regon register per 10 thousand population
Firms, new	Entities of the economy entered, entities of the national economy-indicators, entities indicators, new entities of the national economy recorded in the Regon register per 10 thousand population
Firms with foreign capital participation	Entities of the economy entered, entities with foreign capital participation, entities with foreign capital per 10 thousand population
Firms, domestic	Firms all - firms with foreign capital participation
Price index	Prices, Price indices, Price indices of consumer goods and services, total
Population	Population, Population, Population by domicile/residence and sex, total locations, actual place of residence, as of 31 December, males and females
Employed persons	Labour market, economic activity of the population (average annual data), employed persons by economic sectors and sex, all sectors, total
GDP	Regional accounts, Gross domestic product in current prices, PKD 2007, NACE Rev.2, ESA 1995, Gross domestic product total
Urbanization economies	Population, population, Population by domicile/residence and sex, in urban areas, actual place of residence, as of 31st December, total, divided by population
Industry agglomeration economies	Labour market, Economic activity of the population (average annual data), Percentage of employed persons by economic sectors and sex, industry sector, total
Services agglomeration economies	Labour market, Economic activity of the population (average annual data), Percentage of employed persons by economic sectors and sex, services sector, total
Wages	Wages and salaries and social security benefits, wages and salaries, average monthly gross wages and salaries, total
Students	Higher education, higher education institutions, higher education institutions by type, total institutions, students, divided by population
Unemployment	Labour market, Economic activity of the population (average annual data), unemployment rate by place of residence, total
Telephone lines	Transport and communication, Communication, Main telephone lines of all operators, in total, total locations, total subscribers, divided by population
Road network	Transport and communication, Public roads, roads-indicators, roads ward surface per 100 km²
Railway lines	Transport and communication, Rail transport, railway lines-indicators, railway lines standard gauge per 100 km²
Area	Territorial division, Geodetic area, Area, total in km², as of 2009
R&D agglomeration economies	Science and technology, Research and development activity, Employment in R&D-indices, total, divided by employed persons
Patents	Patents applications to the EPO by priority year (pat_ep_rtot), per million population, taken from Eurostat
German-Polish (GE-PL) border	Distance in regard of one of three main car routes from the regional capital cities to the German border cities Penkun, Gorlitz or Frankfurt Oder, in km, information collected from Google maps
EU membership	EU membership dummy which counts 1 in 2004 and the years thereafter

Note: Data were extracted from the Central Polish Statistical Office if not otherwise indicated.

TABLE A2
Summary statistics

Variable		Mean	Std. Dev.	Min	Max	Observations
Firms, all	overall	230507.4	142212.4	87409.23	681032.4	N=128
	between		145844.7	90752.57	622165.8	n=16
	within		11228.51	184057.2	289374	T=8
Firms, new	overall	18831.39	11615.33	5047.349	66365.11	N=128
	between		11257.86	6418.261	47714.76	n=16
	within		3893.875	9197.005	37481.74	T=8
Firms with foreign	overall	1179.856	1527.748	84.35819	8585.327	N=128
capital	between		1527.23	113.7235	6460.619	n=16
participation	within		360.7532	-350.4603	3304.564	T=8
Firms, domestic	overall	229327.6	140853.7	86986.96	672447.1	N=128
	between		144466.9	90638.85	615705.2	n=16
	within		10928.11	184407.6	286069.5	T=8
GDP	overall	63064.16	51283.44	18030	273612.5	N=128
	between		50979.45	22634.02	215692.7	n=16
	within		13203.56	15114.49	120984	T=8
Urbanization	overall	59.46227	10.02436	40.35843	79.0194	N=128
economies	between		10.30746	40.60549	78.57063	n=16
	within		.3161661	58.54808	60.33716	T=8
Industry	overall	29.71719	5.326019	18.1	40.7	N=128
agglomeration	between		5.190725	20.075	39.4625	n=16
economies	within		1.705266	26.17969	34.59219	T=8
Services	overall	52.80703	5.965454	41	65.1	N=128
agglomeration	between		5.891454	44.3125	62.4125	n=16
economies	within		1.670485	48.19453	57.55703	T=8
Wages	overall	2381.69	394.1925	1875.41	3890.47	N=128
	between		287.1983	2131.069	3302.779	n=16
	within		278.3006	1972.742	2969.381	T=8
Students	overall	4.658047	.907072	2.79	6.8	N=128
	between		.8940095	3.55	6.6825	n=16
	within		.2599616	3.764297	5.178047	T=8
Unemployment	overall	14.78281	5.724421	5.5	26.3	N=128
	between		2.077042	12	18.6	n=16
	within		5.356555	4.657812	24.45781	T=8
Telephone lines	overall	27.93209	4.886204	18.27626	39.10935	N=128
	between		3.18446	22.87786	34.95792	n=16
	within		3.780627	19.95608	34.43345	T=8
Road network	overall	85.71094	30.47313	50.5	170.6	N=128
	between		31.25394	51.2375	164.3875	n=16
	within		2.375445	77.69844	94.24843	T=8
Railway lines	overall	6.920313	3.16582	3.3	18.9	N=128
	between		3.249849	3.5625	17.675	n=16
	within		.2074403	6.445313	8.145312	T=8
Area	overall	19542.44	6645.295	9412	35558	N=128
	between		6836.369	9412	35558	n=16
	within		0	19542.44	19542.44	T=8
R&D	overall	.4190625	.25207	.14	1.32	N=128
agglomeration	between		.2521117	.155	1.16875	n=16
economies	within		.0590109	.2303125	.6015625	T=8
Patents	overall	3.828463	2.959698	.091	14.383	N=123
	between		1.927724	.8206667	7.444625	n=16
	within		2.294847	-.8627867	12.88221	T-bar=7.6875
GE-PL border	overall	371.3187	188.1848	29.5	686	N=128
	between		193.5957	29.5	686	n=16
	within		0	371.3187	371.3187	T=8
EU membership	overall	.75	.4347141	0	1	N=128
	between		0	.75	.75	n=16
	within		.4347141	0	1	T=8

FOOTNOTES

[1]A comprehensive survey on micro-founded analyses of agglomeration economies is offered by Quigley (1998).

[2]See Arauzo-Carod et al. (2010) for a summary of recent empirical contributions.

[3]Cieślik (2005a) observes, however, that this type of agglomeration externalities might have a centrifugal content, considering that strong urbanization dynamics lead to intensified congestion and pollution pressure.

[4]The full list of variables with their descriptions is contained in Table A1 of the Appendix. Table A2 of the Appendix shows basic summary statistics.

[5]All variables that are denoted in values were converted into constant 2002 prices using the regional consumer price index for goods and services.

[6]Also, the relative number of patents as a measure of the actual knowledge base is only an approximation and thus as an imperfect measure of the regional degree of innovativeness. Indeed, not every innovation ends up being patented, and a considerable stock of locally available knowledge remains unprotected and constitutes an even more important source of knowledge spillovers. The difficulty of grasping this phenomenon by means of a single variable is non-negligible.

[7]In 2011, the total volume of investment of businesses in Mazowieckie amounted to 29687.6 mln zloty and constituted 22.7% of investment made in Poland (GUS 2013).

[8]For all regression models, the negative binomial estimator proved to be preferable over the Poisson estimator, as can be seen by the alpha test results in the following regression output tables.

[9]Note that in 1999 a reform in the regional classification system took place that exchanged the formerly 49 voivodeships for 16 voivodeships, now making up the NUTS II level data.

[10]Accession to the EU implied the application of the four principal freedoms of movement of goods and services and of production factors, although the free movement of labor has been restricted by the majority of the old EU members (specifically, Austria, Denmark, Finland, Germany, Greece, Italy, Netherlands, Portugal and Spain) for an initial period of a maximum of seven years.