

Fuchs, Michaela

Conference Paper

How important are agglomeration effects for plant performance? Empirical evidence for Germany

51st Congress of the European Regional Science Association: "New Challenges for European Regions and Urban Areas in a Globalised World", 30 August - 3 September 2011, Barcelona, Spain

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Fuchs, Michaela (2011) : How important are agglomeration effects for plant performance? Empirical evidence for Germany, 51st Congress of the European Regional Science Association: "New Challenges for European Regions and Urban Areas in a Globalised World", 30 August - 3 September 2011, Barcelona, Spain, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<http://hdl.handle.net/10419/120125>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

How important are agglomeration effects for plant performance? Empirical evidence for Germany

Michaela Fuchs¹ Oliver Ludewig²

Abstract

The question whether agglomeration effects are of importance for regional development has a long tradition in regional science. This paper asks if regional characteristics and specifically agglomeration effects influence the performance of plants in Germany and, if so, in which direction. Hence, we aim at contributing to the still sparse empirical studies in this field of research by adding three aspects to the existing evidence. First, we provide the first plant-level evidence on regional agglomeration effects for Germany. Second, while earlier papers looked only at few sectors of the economy or only at manufacturing, we extend our analysis to the services sector. Finally, we are among the first who identify agglomeration effects while controlling for the internal structure of the establishments using a rich set of plant characteristics that are likely to influence productivity. To this end we estimate plant-level production functions augmented by regional characteristics and controlling in detail for plant-specific features. The analysis is conducted both within a static and a dynamic panel framework. We use the IAB Establishment Panel, a large-scale German establishment survey covering around 16,000 establishments each year. In the static framework we find support for the positive impact of localization economies and market size, while urbanization seems to have a negative influence. Results for the dynamic models are rather inconclusive.

Keywords: firms, productivity, externalities, agglomerations

JEL classifications: R30, R12, L25

preliminary version
March 3, 2010

¹ Institute for Employment Research, Regional Research Network Saxony-Anhalt/Thuringia, email: michaela.fuchs@iab.de

² Institute for Employment Research, Regional Research Network, email: oliver.ludewig@iab.de

1. Introduction

Agglomeration effects on regional development and productivity are one major issue in regional science. The idea that regional characteristics influence the productivity of firms that are located in the respective region dates back as far as to Marshall (1890). Knowledge and information spillovers, labor pooling as well as backward and forward linkages give in this view rise to productivity increases. There is a vivid discussion whether these positive externalities arise from specialization or diversification, like Jacobs (1969) postulated. In order to detect agglomeration effects many studies in the field of regional science were dedicated to these issues.

Beugelsdijk (2007) and Raspe/van Oort (2008) argue that these relationships should most profoundly hold at the micro or firm level. Between single plants within a sector or region there is considerable heterogeneity, and within each plant complex processes of employment, output and productivity growth interact (Raspe/van Oort 2008, 104). However, most of the existing papers look at regional aggregates (e.g. Glaeser et al. 1992 or Henderson et al. 1995) while the theoretical arguments focus on the establishment level.

Only since Henderson's (2003) seminal paper there is an increasing effort to look into these relationships at the micro level that is at the level of firms, establishments or plants. This is surprising because one should expect that the agglomeration externalities would show primarily on the plant production function (Cingano/Schivardi 2004, Rosenthal/Strange 2004). Hence, in this study we ask: do regional agglomeration characteristics influence establishment productivity and, if so, in which direction?

This question has been asked in a rising number of papers. Most of them focus on the manufacturing sector or even sub-sectors (e.g. Henderson 2003, Saito/Gopinath 2009, Anderson/Löf 2009, Baldwin et al. 2010), the Anglo-Saxon countries (e.g. Henderson 2003, Maré/Graham 2010) and use administrative data (e.g. Maré/Graham 2010, Henderson 2003, Saito/Gopinath 2009, Baldwin et al. 2010, Cingano/Schivardi 2004), which allows little insight in the organization of the plant. Summarizing this literature there is strong focus on the manufacturing sector and the Anglo-Saxon countries. Evidence for the whole range of sectors is rare and other countries than the Anglo-Saxon ones are seldom investigated. Specifically there is to our knowledge no study analyzing agglomeration effects on the establishment level for Germany. The only recent paper looking at productivity with German data uses regionally aggregated data (Eckey et al. 2010). In addition, there is a huge body of literature showing the importance of the internal structure of an establishment, e.g. industrial relations, part-time work or the owner structure, for its productivity (e.g. Ichniowsky et al. 1997, Black/Lynch 2001, Ludewig/Sadowski 2009). However, these elements are missing in studies looking at agglomeration effects on plant level.

We want to contribute to this research area by adding three aspects to the existing evidence. First, we will provide the first plant-level evidence on regional agglomeration effects for Germany. Second, while earlier papers looked only at few sectors of the economy or only at manufacturing, we also include the services sector. Finally, we will be among the first who identify the agglomeration effects while controlling the internal structure of the establishment using a rich set of plant characteristics that are likely to influence productivity. Additionally, we use for the production factor labor the number of

hours worked instead of the usual number of employees per plant. This gives a more precise account of the labor input.

In order to quantify the impact of geography on plant performance we estimate plant-level production functions augmented by regional characteristics and controlling in detail for plant-specific features. The analysis is conducted both in a static and a dynamic panel framework. The static model answers the question if the output level of a plant depends on the regional economic structure, controlling for plant characteristics. The dynamic model deals with the endogeneity problems that can evolve in the static models and focuses on the temporal persistence of a plant's performance subject to the surrounding geographic features.

The analysis is based on the IAB Establishment Panel. It is an annual representative employer survey at individual establishments in Germany covering a broad range of establishment characteristics. We use the years from 2001 to 2006 with roughly 20,000 observations for almost 4,000 establishments. The regional variables are added on the NUTS3-level.

The paper is structured as follows. Section 2 discusses the theoretical background and the econometric setup. In section 3, the data and the variables are described. Section 4 is dedicated to the results, and section 5 concludes.

2. Theoretical background and econometric setup

2.1 Theoretical background of agglomeration economies and productivity

The literature usually distinguishes two types of positive agglomeration effects: localization effects and urbanization effects. The term localization effects subsumes those effects which arise from the spatial concentration of companies of the same industry (Rosenthal/Strange 2004, Eckey et al. 2010). Already Marshall (1890) identified three different channels of these localisation effects. Labor pooling allows better matches, reduces search costs, training costs and risks for both employer and employee (Rosenthal/Strange 2004). Knowledge spillovers induce product and process innovations and thus increase productivity. In the localization framework spillovers especially of tacit knowledge are the more likely the more employees of the same industry interact with each other and switch employers (Jaffe et al. 1993, Rosenthal Strange 2004, Baldwin et al. 2010). The third effect stems from forward and backward linkages. The common usage of an input or a supplier can for example free up economies of scale, of which benefit all establishments in that specific region and industry (Baldwin et al. 2010, Andersson/Lööf 2009).

Urbanization effects instead arise from the spatial concentration of establishments from different industries. Diversity facilitates the development of new ideas through cross-fertilization and allows for knowledge spillovers across industries (Jacobs 1986, Rosenthal/Strange 2004, Frenken et al. 2007, Baldwin et al. 2010). Additionally the New Economic Geography emphasizes the importance of a big home market for producers with substantial transport costs (Krugman 1991).

Apart from these agglomeration effects there are “deglomeration” – e.g. congestion – effects working in the opposite direction. For example, in densely populated areas the overcrowding of places has un-

favourable consequences: increasing land prices, traffic problems, environmental pollution, and an overstrained infrastructure etc. increase the cost of production and thus reduce productivity (Eckey et al. 2010).

The question whether there are positive agglomeration effects and if localization or urbanization effects dominate is an old one as the cited literature indicates. In the course of time many empirical studies have attempted to answer these questions. However, a huge part of them is based on regional aggregates (e.g. Glaeser et al. 1992, Henderson et al. 1995, Fuchs 2011). While such studies found evidence for positive agglomeration effects they look at the wrong level of analysis (Rosenthal/Strange 2004, Cingano/Schivardi 2004). Localization and urbanization effects work on the micro level. They shift the production function of establishments (Rosenthal/Strange 2004).

A still small but rising number of papers that try to identify agglomeration effects on the plant level has been published recently. Henderson (2003), for example, finds in his paper for four high-tech and five machinery goods industries in the USA that there are substantial localization economies in the high-tech industries. Urbanization economies seem to dominate the machinery good industries. There is only weak evidence for dynamic effects. Andersson and Löff (2009) identify substantial agglomeration effects for the Swedish manufacturing sector but they do not differentiate between localization and urbanization effects. Baldwin et al. (2010) find strong evidence for localization effects in the Canadian manufacturing sector. Cingano and Schivardi's (2004) results indicate for an Italian firm panel that there are substantial localization effects. Maré and Graham (2010) analyze agglomeration effects in New Zealand on the one-digit industry level using a huge firm level panel data set. Using employment density, they find substantial agglomeration elasticities, however they do not differentiate between localization and urbanization effects. Moretti (2004) shows that productivity is higher in regions with higher shares of college graduates. Summing up, the available evidence is more in favor of the localization effects but on balance it seems still somehow inclusive.

For Germany there is to our knowledge until yet no study that has analyzed the relationship between regional size and firm performance with firm-level data and conditioned on attributes of firms. One notable exception is Eckey et al. (2009), but they resort to a regional production function in order to estimate the influence of agglomeration and population effects on the economic activity of spatial units. Hence, we estimate in the following sections establishment-level production functions including detailed measures of regional features that account for regional concentration, specialization and diversity.

2.2 Econometric setup

The starting point of our empirical model is a Cobb-Douglas establishment-level production function for plant i with the stock of capital and labor included as inputs (see also Cingano/Schivardi 2004). Henderson (2003) and Moretti (2004) report little differences in parameter estimation between Cobb-Douglas or more flexible forms of production function. Our basic model is

$$Y_{it} = AK_{it}^{\alpha}L_{it}^{\beta}, \quad (1)$$

with the subscript $i=1, 2, \dots, N$ referring to a plant and $t=1, 2, \dots, T$ to a point in time. Y_{it} denotes value added, A the technology that is available for all plants, K_{it} is plant i 's capital stock and L_{it} expresses the amount of labor employed in plant i . To obtain a linear expression and to account for random measurement errors and stochastic shocks, we take the natural logarithm and add the random error term ε_{it} ,

$$\ln Y_{it} = \ln A + \alpha \ln K_{it} + \beta \ln L_{it} + \ln \varepsilon_{it}. \quad (2)$$

Equation (2) is our basic estimation equation that explains plant output by the two factors of production. Throughout the analysis, we assume that other characteristics of the firm and the characteristics of the location of the firm are transmitted through the multifactor productivity term A . Similar modeling frameworks are applied by Henderson (2003), Moretti (2004), Andersson/Lööf (2009) and Baldwin et al (2010). This results in the following specification of A :

$$\ln A = \gamma + \varphi \ln X_{it} + \vartheta \ln G_{rt} + \lambda_{rt} + \eta_{it}, \quad (3)$$

where X is a vector of characteristics related to plant i and G is a vector of characteristics that are associated with region r . In this formulation, agglomeration phenomena ‘shift’ the production function of the establishments (see Rosenthal/Strange 2004). Substituting equation (3) into equation (2), we obtain our full model:

$$\ln Y_{it} = \gamma + \varphi \ln X_{it} + \vartheta \ln G_{rt} + \lambda_{rt} + \eta_{it} + \alpha \ln K_{it} + \beta \ln L_{it} + \ln \varepsilon_{it} \quad (4)$$

In containing variables for both the plant and the region, model (4) combines information on two different levels of observation, with some of them not varying between plants or regions. This multilevel structure can result in inefficient estimates of the coefficients and in biased estimates of the standard errors especially of the variables for the higher level (Moulton 1990). In order to deal with this problem clustering-robust linear regression techniques are used to estimate standard errors that recognize this clustering of the data. This method relaxes the independence assumption and requires only that the plant-level observations be independent across regions. By allowing any given amount of correlation within regions, clustering-robust techniques estimate appropriate standard errors when many observations share the same value on some but not all independent variables.

For estimating equation (4), we focus on methods for a short panel, meaning data on many plants but few time periods. In a first step, we run static panel regressions based on the OLS, fixed effects (FE) and random effects (RE) estimation techniques. One of the main benefits of the FE estimator is its ability to control for time-invariant unobserved heterogeneity. This benefit comes with downsides, however. The first downside is that there are no estimates for time invariant variables. The FE estima-

tors use only the within variance and forgo the between variance (Baltagi 2001). This concentration on the within variance results in the second downside: coefficient estimates for variables with low within variance are inefficient (Plümper/Tröger 2007).

However, most regional variables show a substantial inertia and change only slowly over time, and as a consequence their within variance is small compared to the between variance (e.g. Anderson/Lööf 2009). This is at least true for the regional variables we use in this paper. Plümper/Tröger (2007) provide their FE vector decomposition estimator as a solution (see for an application Anderson/Lööf 2009). However, Greene (2010) and Breusch et al. (2010) show severe problems of this approach. Thus, we rely on an older two-step solution, which was already applied within a production function framework. In this two-step approach (see for example Black/Lynch 2001), a common FE estimator is used to estimate a baseline production function building on the variables which vary substantially across time. Then the fixed effect is identified and used as dependent variable. In the second step this establishment-specific effect is explained by time-invariant variables and variables with only small within variance. Although the within variance for these quasi-fixed variables is small, it is still greater than zero. Hence, we take the mean of these variables across time for each establishment and then we regress for one observation for each establishment the fixed effect on these means.

In the second step of our econometric analysis, we resort to dynamic panel data models. The inclusion of a lagged dependent variable $X_{i,t-m}$ dating back over an a priori unknown period of time ($m=1, \dots, M$) allows for the explicit modeling of adjustment processes and the influence of history on plant output.

$$\ln Y_{it} = \gamma + \varphi \ln X_{it} + \sum_{m=1}^M \delta_m Y_{i,t-m} + \vartheta \ln G_{rt} + \lambda_{rt} + \eta_{it} + \alpha \ln K_{it} + \beta \ln L_{it} + \ln \varepsilon_{it} \quad (5)$$

In dynamic panel models such as (5), OLS estimates are biased and inconsistent if the lagged dependent variable is correlated with the error term, as it is frequent in dynamic panels with a short time dimension (Nickell 1981). We therefore estimate equation (5) with a generalized methods of moments (GMM) estimator and use the two-step system GMM estimator proposed by Arellano/Bover (1995) and Blundell/Bond (1998). In comparison to the difference GMM estimation technique of Arellano/Bond (1991) it allows the introduction of more instruments, which improves efficiency.³ One further advantage that is important for our analysis is that, in contrast to the difference GMM estimator, it can include time-invariant regressors.

³ Since the Arellano/Bond estimator is based only on an equation in differences, it is also called difference GMM estimator as opposed to the system GMM estimator (see, for example, Roodman 2009).

3. Data and variables

3.1 Data

For information on the level of the individual plants we resort to the IAB Establishment Panel, an annual representative employer survey at individual establishments in Germany (for details see Fischer et al. 2009). Approximately 16,000 establishments from all sectors of the economy and of all sizes are questioned on a large number of employment-related subjects, including employment development, business policy and development, innovations, wages and salaries, working times and general data on the establishment. The Establishment Panel was started in Western Germany in 1993 and in Eastern Germany in 1996. As a comprehensive longitudinal data set, it forms the basis for detailed research into the determinants of plant performance.⁴

We consider only those sectors that are subject to market-based forces, i.e. where our assumptions of a cost-minimizing firm are appropriate. Hence, plants that belong to the public sector are excluded. In addition, those plants are excluded that express their business volume by the budget volume (administration and property budget), i. e. publicly owned establishments that belong to market-oriented sectors. Sectors that are strongly dependent on geographical features (agriculture, fishing and mining) are not considered either. Finally, we exclude establishments which have missing values for our relevant variables. The period of observation covers the years from 2001 to 2006. Our final panel data set comprises a total of 20,106 observations on 4,891 plants.

The variables characterizing the plant's environment are calculated at the NUTS-3 level that comprises 439 *Kreise* and *kreisfreie Städte*. Table 1 lists all the data sources for the regional variables used.

3.2 Variables

For the estimation of the basic production function (2), information on output, capital, and labor are needed. We measure the dependent variable of a plant's **output** with its value added in the respective year. Regarding the input of **capital**, the IAB Establishment Panel provides no direct information. As Müller (2008, 358) notes, this is a general problem of many establishment data sets. Measurement errors in capital stock, however, will lead to biased estimates and any inference based on such estimates could be misleading. In order to get a reliable measure of the capital stock, we calculate it according to a modified perpetual inventory approach specifically developed for this purpose by Müller (2008). The input of **labor** is measured by the number of hours worked by the full-time equivalents in an establishment. The IAB Establishment Panel provides us with detailed information on the number of hours worked for full-time and for part-time workers. First of all, the establishments are asked about the agreed working hours per week for full-time employees at present. As regards the part-time employees, the establishments are asked to group them according to working less than 15 hours per week, between 15 and 25 hours or more than 25 hours. Taking the total number of hours per week as refer-

⁴ English versions of the questionnaire can be downloaded under http://fdz.iab.de/en/FDZ_Establishment_Data/IAB_Establishment_Panel/IAB_Establishment_Panel_Working_Tools.aspx

ence, we can then calculate the number of full-time equivalents and the respective number of hours worked per establishment.

Region-specific variables

Following the literature on agglomeration economies discussed in section 2.1, three types of externalities acting as centripetal forces of agglomeration can be distinguished. First, localization economies emerge if plants of the same industry are located close to each other. Hence, they should be fostered by a specialized economic structure of a region. Second, urbanization economies arise if plants belonging to different industries are spatially close, which should be advanced by a diversified economic structure. Third, the general features of large markets with a greater demand make it possible for plants to adopt more efficient production processes. We attempt to quantify the impact of each of these externalities with specific measures of the regional economic structure and add further more general regional indicators. The respective indices used in this study are calculated on the basis of all employees liable to social security that are provided by the Federal Employment Office.

The degree of specialization in a region that can give rise to **localization externalities** is measured with the Krugman specialization index (Südekum 2006):

$$KSI_z = \sum_s \left(\left| \frac{L_{zs}}{L_z} - \frac{L_s}{L} \right| \right).$$

It corresponds to the absolute value of the difference between the share of employment L in region z and sector s on total employment in region z and the corresponding share on the national level. The values for KSI range between zero and two. If KSI is equal to zero, the region under consideration has the same economic structure as the national average. A value of two indicates that there is no sector that exists in both regions simultaneously.

Another way to measure localization externalities is via the geographic concentration of economic activities. Establishments that form part of a spatially concentrated sector can profit from positive agglomeration externalities at work in this sector (Holmes/Stevens 2002, Südekum 2006). We measure the spatial **concentration** of sectors with a localization quotient that is calculated at the 3-digit industry level (see also O'Donoghue/Gleaves 2004):

$$LQ_{zs} = \frac{L_{zs}/L_s}{L_z/L}$$

The share of employment L in region z and sector s is divided by the share of employment in sector s on the national level. If LQ is smaller than one, the sector under consideration is represented in region z below average. Values larger than one indicate that the sector is concentrated above average.

We measure the degree of economic diversity that can give rise to **urbanization externalities** with a Hirshman-Herfindahl index across the number of sectors per region (Combes 2000, Combes et al 2004, Mameli et al 2008):

$$DIV_z = \left[\sum_{s=1}^I \left(\frac{L_{zs}}{L_z} \right)^2 \right].$$

DIV_z is zero if local employment is concentrated in only one sector and equals the logarithm of the number of sectors if employment is distributed uniformly across sectors. If urbanization externalities should exist in a region, we would expect DIV_z to have a positive influence on plant productivity.

For measuring **market size**, we use GDP per district as a proxy for general agglomeration effects that are related to the size of a region. One advantage of resorting to GDP instead of population is that GDP is measured exactly in the region where it is produced, while people do not necessarily live where they actually work.

One last indicator relevant for the economic prospects of a region is its **accessibility**. It covers the geographical position as well as the existence of a good road and rail transport system, an airport or a harbor. The accessibility of a region decisively influences the costs for the transport of goods and hence its integration with other regions (see also Ottaviano/Puga 1998 and Hoogstra/van Dijk 2004). It is calculated as the average driving time in minutes by car to the nearest highway entry.

Plant-specific control variables

In addition to the region-specific characteristics, we control for several plant-specific variables that are expected to influence plant performance.

Engaging in **exports** is an important opportunity for a plant to expand the market for its products. On the other hand, the establishment is relatively prone to negative demand shocks emanating from the export partners, so that a priori there is no clear-cut relationship. Exports also imply a higher degree of competition, because exporters are selling in additional markets with additional competitors. Competition puts pressure on the establishment to be more efficient and thus exporters should be more productive (see Wagner 2007). We use the share of sales achieved abroad on total sales in the last fiscal year as indicator of export orientation.

Plant productivity can also be indirectly influenced by the existence of a **works council** that has a strong impact on many decisions of the plant and the organisational structure (see, e.g., Blanchflower et al 1991). Additionally, empirical studies find a positive correlation between productivity and works councils (e.g. Hübler/Jirjahn 2003 and for an overview Hübler 2003), thus we have to control for it. This is captured by a dummy variable that take the value of one if the establishment has a works or

staff council or some other company-specific form of staff representation (staff spokesperson, round table conferences).

The influence of external enterprises can generate both positive and negative productivity effects (Bates 1995). On the one hand, the provision of technological or entrepreneurial know-how or networks with customers or suppliers can stimulate positive productivity effects. On the other hand, they can also be negative if an establishment is strongly influenced by strategies and decisions of the external partner. The IAB-Establishment Panel asks if the establishment surveyed is a) an **independent** company or an independent organization without other places of business, b) the head office of an enterprise or an organization with other places of business/offices/branches, c) a place of business/office/branch of a larger enterprise or organization.⁵ The answers are captured with the help of dummy variables. We also include a dummy variable if the enterprise is mainly or exclusively in foreign **property** (see also Henderson 2003 and Baldwin et al 2008).

Plant performance also strongly depends on plant age. One can argue that, over time, plants age and slowly lose their ability to compete. On the other hand, plants can also improve with age due to learning effects (for an overview see Coad et al 2010). Hence, a priori the influence of age on plant performance is important, but if it is positive or negative needs to be answered empirically. We measure plant age with a dummy variable that takes on the value of 1 if the plant is young, i.e. if it was founded after 2001.

Last, productivity is expected to vary significantly between **sectors**. We control for sector-specific effects by introducing sectoral dummies at the 2-digit level of the WZ 2003 (corresponding to the NACE Rev.1 classification). Since systematic differences between Eastern and Western Germany still exist, we also include a dummy variable for the two parts of the country. Table 2 summarizes all variables under consideration. Descriptive statistics are provided in table 3.

4. Results

4.1 Static model

In this section, we concentrate on the results of the static specification and estimate three models in total. In order to quantify the basic relationship between the two factors of production and plant output, we first estimate exclusively the influence of K_{it} and L_{it} on Y_{it} (model bl). Then, as a consistency check we add the plant-level control variables (model cl), and finally we move to the full model (4) that includes the regional variables (model reg):

⁵ It is also asked if it is a middle-level authority of a multi-level company or a multi-level authority/organization. But since this includes basically the public sector, these enterprises are a priori excluded from the analysis.

$$\ln Y_{it} = \gamma + \varphi \ln X_{it} + \vartheta \ln G_{rt} + \lambda_{rt} + \eta_{it} + \alpha \ln K_{it} + \beta \ln L_{it} + \ln \varepsilon_{it} \quad (4)$$

Since our interest is on the impact of the regional variables, in the following discussion of the results we do not go into the plant characteristics in details. Table 4 presents the results of the static model. We start off with a simple pooled OLS regression in columns (1) to (3), present the results of the RE and the FE estimations in columns (4) to (9) as well as the results of the two-step method in columns (10) to (12).

A comparison of the four estimation methods shows that the OLS, the RE and the two-step results are broadly comparable in the significance and in the size of the estimates. The FE estimator yields diverging results, which is not surprising given that most of the regional variables show a substantial inertia over time.

All three consistent estimation methods emphasize the highly significant and positive role of labor and capital in explaining output even after including the plant and regional variables. Hence, the underlying fundamental relationships seem to be robust. Among the regional characteristics, the degree of specialization of the regional economy is not significant. On the other side, the location of a plant in a region where the sector the plant belongs to is concentrated geographically exerts a positive and highly significant impact on its performance. This implies that, obviously, localization economies are important as long as they act not through the specialized economic structure of the region but through the concentration of the sector the plant belongs to in the respective region.

The estimates on the degree of diversity are significant and negative. This result can be interpreted in the sense that urbanization economies that are transmitted through a diversified economic structure are not reflected in plant output. On the contrary, the fewer sectors there are in a region, the higher is output.

GDP density turns out to be highly significant and positive. This can be taken as evidence for the existence of global agglomeration economies that come into force in large markets. This result is also in line with the findings of Andersson/Lööf (2009). The last regional indicator, in contrast, has a rather weak impact on plant output.

sectoral regressions to come.....

To sum up, the first results from the static regressions provide evidence of a strong impact of regional characteristics on plant performance. Especially the geographic concentration of plants from the same industry and the market potential act as promoting measures, what can be interpreted as evidence for the role of localization and global agglomeration economies. Urbanization economies, on the other hand, do not seem to be effective.

Table 4: Production function estimates using OLS, RE, FE and the two-step approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS_bl	OLS_cl	OLS_reg	RE_bl	RE_cl	RE_reg	FE_bl	FE_cl	FE_reg	zs_bl	zs_cl	zs_reg
ln l	0.915*** (95.80)	0.842*** (73.38)	0.832*** (71.93)	0.896*** (67.79)	0.831*** (55.60)	0.821*** (53.97)	0.506*** (12.91)	0.487*** (12.35)	0.483*** (12.15)	0.506*** (12.91)		
ln k	0.190*** (25.80)	0.165*** (20.08)	0.167*** (20.05)	0.199*** (19.63)	0.180*** (16.21)	0.182*** (16.27)	0.161*** (3.95)	0.133*** (3.01)	0.135*** (3.04)	0.161*** (3.95)		
ln spec			-0.047 (-0.65)			-0.082 (-0.71)			0.180 (0.45)			0.147 (1.10)
ln conc			0.028** (3.19)			0.033** (2.32)			0.069* (1.86)			0.136*** (7.74)
ln div			-0.094** (-2.15)			-0.078 (-1.20)			0.415** (2.06)			-0.190** (-2.44)
ln gdp area access			0.048*** (5.92)			0.049*** (3.80)			-0.141 (-0.59)			0.079*** (5.38)
			0.002** (2.10)			0.002 (1.29)			0.001*** (2.64)			0.003* (1.89)
export		0.004*** (8.19)	0.004*** (7.74)		0.003*** (5.26)	0.003*** (5.04)		0.001 (1.11)	0.001 (1.03)		0.007*** (8.29)	0.006*** (7.18)
counc		0.206*** (8.30)	0.202*** (8.16)		0.157*** (4.65)	0.151*** (4.47)		-0.086* (-1.66)	-0.083 (-1.59)		0.869*** (23.04)	0.787*** (20.50)
struct2		0.238*** (8.24)	0.233*** (8.07)		0.168*** (4.92)	0.169*** (4.92)		0.025 (0.53)	0.034 (0.71)		0.331*** (6.31)	0.323*** (6.16)
struct3		0.099*** (3.24)	0.092*** (2.99)		0.113*** (3.45)	0.109*** (3.33)		0.091** (2.33)	0.094** (2.40)		0.398*** (6.96)	0.353*** (6.26)
fprop		0.032 (0.90)	0.024 (0.66)		0.048 (1.02)	0.045 (0.95)		-0.091 (-1.29)	-0.085 (-1.19)		0.121* (1.81)	0.097 (1.48)
young		-0.060 (-1.28)	-0.063 (-1.33)		-0.033 (-0.62)	-0.031 (-0.57)		0.040 (0.42)	0.054 (0.56)		-0.114 (-1.53)	-0.110 (-1.47)
east	- 0.132*** (-7.54)	-0.131*** (-7.48)	-0.091*** (-4.84)	- 0.139*** (-4.55)	- 0.142*** (-4.73)	- 0.098*** (-3.02)					- 0.230*** (-7.11)	- 0.170*** (-5.00)
_cons	4.847*** (74.30)	5.402*** (57.96)	4.966*** (29.25)	4.822*** (48.39)	5.256*** (39.08)	4.851*** (19.58)	8.207*** (13.14)	8.777*** (13.22)	-20.996* (-1.76)	8.207*** (13.14)	- 1.035*** (-5.58)	- 1.377*** (-5.15)
N	20106	20106	20046	20106	20106	20046	20106	20106	20046	20106	4594	4592
R ²	0.710	0.728	0.729				0.013	0.015	0.016	0.013	0.338	0.353
adj. R ²	0.710	0.727	0.728								0.332	0.346
pseudo R ²												
N_g				4891.00 0	4891.00 0	4887.00 0	4891.00 0	4891.00 0	4887.00 0	4891.00 0		
r2_w				0.013	0.013	0.014	0.013	0.015	0.016	0.013		
F	1.5e+04	1202.442	1083.881				98.351	16.720	13.086	98.351		
chi2				1.7e+04	1.9e+04	2.0e+04						

t statistics in parentheses. Time and sector dummies included but not reported.

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

4.2 Dynamic model

This section presents results of the dynamic equation (5). Just as in the case of the static models, we first estimate the basic model (bl) and successively add the plant (model cl) and regional characteristics (model reg). The models are estimated with the system GMM estimator proposed by Arellano/Bover (1995) and Blundell/Bond (1998), and we report heteroskedasticity-robust standard errors calculated according to the mechanism by Windmeijer (2005). After trying out several specifications for the number of lags included, our optimal specification of equation (5) includes only one lag of the dependent variable:

$$\ln Y_{it} = \gamma + \varphi \ln X_{it} + \delta \ln Y_{t-1} + \vartheta \ln G_{rt} + \lambda_{rt} + \eta_{it} + \alpha \ln K_{it} + \beta \ln L_{it} + \ln \varepsilon_{it} \quad (6)$$

Before turning to the results, the validity of the system GMM estimator should be checked. A comparison of the regression results for the lagged endogenous variable with the system GMM estimator on the one hand and the OLS and FE estimator on the other hand can serve as an indicator of its validity (Bond 2002, Roodman 2009). Bond (2002) notes that the OLS estimation results for equation (6) are biased upwards, whereas the FE estimation results are biased downwards. Accordingly, consistent GMM results should lie between those of the two former estimators.

Table 5 presents the results for the OLS, the FE and system GMM estimation methods and for the three models, respectively. In all cases, the system GMM estimator yields a value for the coefficient on Y_{t-1} that lies between those of the OLS and the FE estimator. Hence, the OLS estimator gives an upper bound and the FE estimator a lower bound for the coefficient estimated with the system GMM technique (see also Andersson/Löf 2009). As a consequence, the system GMM estimator should generate consistent results.

Table 5: Comparison of the results on Y_{t-1}

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS_bl	OLS_cl	OLS_reg	FE_bl	FE_cl	FE_reg	GMM_bl	GMM_cl	GMM_reg
$\ln y_{t-1}$	0.516*** (33.79)	0.482*** (30.75)	0.480*** (30.42)	-0.067*** (-5.70)	-0.068*** (-5.76)	-0.071*** (-6.02)	0.163*** (5.08)	0.158*** (3.39)	0.165** (2.06)
$\ln l$	0.426*** (24.68)	0.420*** (21.54)	0.418*** (21.40)	0.567*** (9.16)	0.539*** (8.65)	0.535*** (8.54)	0.693*** (5.25)	0.628*** (3.54)	0.698*** (4.00)
$\ln k$	0.091*** (10.69)	0.087*** (8.74)	0.088*** (8.75)	0.202*** (4.31)	0.218*** (4.19)	0.215*** (4.13)	0.147** (2.48)	0.137** (1.99)	0.138 (0.98)
N	9783	9783	9758	9783	9783	9758	9783	9783	9758

t statistics in parentheses.

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

Table 6: Dynamic production function estimates using system GMM

	(7) GMM_bl	(8) GMM_cl	(9) GMM_reg
ln y_{t-1}	0.163*** (5.08)	0.158*** (3.39)	0.165** (2.06)
ln l	0.693*** (5.25)	0.628*** (3.54)	0.698*** (4.00)
ln k	0.147** (2.48)	0.137** (1.99)	0.138 (0.98)
ln spec			-0.783 (-0.70)
ln conc			0.316 (0.58)
ln div			0.258 (0.76)
ln gdp area access			0.099 (1.04) 0.000 (0.68)
export		-0.005** (-2.47)	-0.003* (-1.70)
counc		-0.067 (-0.59)	-0.072 (-0.49)
struct2		-0.027 (-0.38)	-0.021 (-0.29)
struct3		0.074 (1.14)	0.046 (0.63)
fprop		0.030 (0.23)	0.060 (0.47)
young		-0.042 (-0.48)	-0.032 (-0.32)
east			
_cons	4.823*** (3.85)	5.684*** (2.92)	5.148 (1.38)
N	9,783	9,783	9,758
R^2			
adj. R^2			
pseudo R^2			
N_g	2320.000	2320.000	2318.000
F			
chi2	83.950	61.823	668.575
r2_w			
arm1	-12.896	-10.144	-6.702
arm2	1.703	1.541	1.250
sargan			

t statistics in parentheses. Time and sector dummies included but not reported.

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

The preliminary results of the dynamic panel regressions are reported in table 6. Beginning with the impact of history, the estimate of the lagged output is highly significant and positive. Evidently, output is very persistent and exhibits a high path dependency: output is higher today when it was already high

in the previous period. Its magnitude changes only slightly when moving from the basic to the full model, so that these findings can be regarded as robust.

While the same holds for labor input, the capital stock turns insignificant when we include the regional variables. In the dynamic setting, the estimates for all regional variables are not significant. It might be the case that the impact we found in the static models is contained in the lagged dependent variable, which requires further investigation.

Sectoral regressions to come....

....

5. Conclusions

In this paper we investigate the importance of regional characteristics for the performance of plants when controlling for various plant-specific determinants. To this end we augment a Cobb-Douglas production function with labor and the capital stock as inputs by characteristics specific to the plant and to its regional environment. In line with the empirical literature on agglomeration economies, the latter is characterized by regional specialization, sectoral concentration, diversity and market potential.

First results from the static estimation models give support for the positive impact of localization economies and market size, while urbanization seems to have a negative influence. The preliminary results for the dynamic models are rather inconclusive. Therefore, further challenges are to investigate why the influence of the regional variables found in the static models do not seem to exist in the dynamic case. Further work also remains with regard to the sectoral aspects. So far, mainly the manufacturing sector has been analyzed. We want to find out if there are differences between manufacturing and services, which is a major advantage of our dataset. Finally, the dummy indicating the location in Eastern Germany is highly significant in the static regressions. Further work should be dedicated to find out if these systematic differences might be due to general differences in the regional structure of the two regions.

References

- Andersson, M. and H. Lööf (2009): Agglomeration and productivity: evidence from firm-level data, *Annals of Regional Science*, DOI 10.1007/s00168-009-0352-1.
- Arellano, M. and S. Bond (1991): Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* 58(2), 227-297.
- Arellano, M. and O. Bover (1995). Another Look at Instrumental Variable Estimation of Error Component Models, *Journal of Econometrics*, 68, 29-51.
- Baldwin, J., M. Brown and D. Rigby (2010): Agglomeration economies: microdata panel estimates from Canadian manufacturing, *Journal of Regional Science* 50(5), 915-934.
- Baltagi, B. H. (2001): *Econometric analysis of panel data*. Chichester, UK, Wiley and Sons.
- Bartelsman, E. and M. Doms (2000): Understanding productivity: lessons from longitudinal microdata. *Journal of Economic Literature*, 38(3), 569-594.
- Black, S. E. and Lynch, L. M. (2001): How to compete: the impact of workplace practices and information on technology and productivity, *Review of Economics and Statistics* 83(3), 434-445.
- Blanchflower, D., N. Milward and A. Oswald (1991): Unionism and employment behavior, *The Economic Journal* 101(407), 815-834.
- Blundell, R and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel data Models. *Journal of Econometrics*, 87(1), 115-43.
- Bond, S. (2002): *Dynamic panel data models: A guide to micro data methods and practice*, CWP 09/2002.
- Breusch, T., Ward, M. B., Nguyen, H. and Kompas, T. (2010): On the fixed-effects vector decomposition, MPRA Paper No. 21452 (http://mpra.ub.uni-muenchen.de/21452/1/MPRA_paper_21452.pdf)
- Cingano, F. and F. Schivardi (2004): Identifying the sources of local productivity growth, *Journal of the European Economic Association* 2(4), 720-742.
- Coad, A., A. Segarra and M. Teruel (2010): Like milk or wine: does firm performance improve with age? *Papers on Economics and Evolution* 1006, MPI Jena.
- Combes, P. (2000): Economic structure and local growth: France 1984-1993, *Journal of Urban Economics* 47(3), 329-255.
- Combes, P., T. Magnac and J.-M. Robin (2004): The dynamics of local employment in France, *Journal of Urban Economics* 56(2), 217-243.
- Eckey, H.-F., R. Kosfeld and N. Muraro (2009): Effects of population development on the economic level of the regions in Germany, *Jahrbuch für Wirtschaftswissenschaften* 60(3), 204-226.
- Fischer, G., F. Janik, D. Müller and A. Schmucker (2009): The IAB Establishment Panel - things users should know, *Schmollers Jahrbuch* 129, 133-148.
- Frenken, K., F. van Oort, and T. Verburg (2007): Related variety, unrelated variety and regional economic growth. *Regional Studies* 41, 685-697.

- Fuchs, M. (2011): The determinants of local employment dynamics in Western Germany, *Empirical Economics* 40(1), 177-203.
- Glaeser, E., H. Kallal, J. Scheinkman and A. Shleifer (1992): Growth in cities, *Journal of Political Economy* 100(6), 1126-1152.
- Greene, W. (2010): Fixed effect vector decomposition: a magical solution to the problem of time invariant variables in fixed effects models? (http://web-docs.stern.nyu.edu/old_web/economics/docs/workingpapers/2010/Greene_Comment-FixedEffects.pdf)
- Henderson, V. (2003): Marshall's scale economies, *Journal of Urban Economics* 53, 1-28.
- Henderson, V., A. Kuncoro and M. Turner (1995): Industrial development in cities, *Journal of Political Economy* 103(5), 1067-1090.
- Holmes, T. and J. Stevens (2004): Geographic concentration and establishment size: Analysis in an alternative economic geography model, *Journal of Economic Geography* 4(3), 227-250.
- Hoogstra, G. and J. van Dijk (2004): Explaining firm employment growth: Does location matter? *Small Business Economics* 22(3/4), 179-192.
- Hübler, O. (2003) Fördern oder behindern Betriebsräte die Unternehmensentwicklung? *Perspektiven der Wirtschaftspolitik* 4, No. 4, pp. 379-397.
- Hübler, O. and Jirjahn, U. (2003): works councils and collective bargaining in Germany: the impact on productivity and wages, *Scottish Journal of Political Economy* 50(4), 471-491.
- Ichniowski, C., Shaw, K. and Prennushi, G. (1997), The effects of human resource management practices on productivity: a study of steel finishing lines, *American Economic Review* 87, 291-313.
- Jaffe, A., M. Trajtenberg, and R. Henderson (1993): Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics* 108, 577-598.
- Jacobs, J. (1986): *Cities and the Wealth of Nations*, Harmondsworth.
- Jacobs, J. (1969). *The Economy of Cities*. New York; Vintage.
- Krugman P (1991) *Geography and trade*. The MIT Press, Cambridge
- Mameli, F., A. Faggian and P. McCann (2008): Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation, *Spatial Economic Analysis* 3(3), 343-360.
- Marshall, A. (1890): *Principles of Economics*. London:Macmillan.
- Moulton, B. (1990): An illustration of a pitfall in estimating the effects of aggregate variables on micro units, *Review of Economics and Statistics* 72(2), 334-338.
- Müller, S. (2008): Capital stock approximation using firm level panel data, *Jahrbücher für Nationalökonomie und Statistik* 228/4, 357-371.
- Nickell, S. (1981): Biases in dynamic models with fixed effects, *Econometrica* 49(6), 1417-1426.
- O'Donoghue, D. and B. Gleave (2004): A note on methods for measuring industrial agglomeration, *Regional Studies* 38(4), 419-427.

- Ottaviano, G. and D. Puga (1998): Agglomeration in the global economy: A survey of the 'New Economic Geography', *World Economy* 21(6), 707-731.
- Plümper, T. and V. Troeger (2007): Efficient estimation of time-invariant and rarely changing variables in finite sample panel analysis with unit fixed effects, *Political Analysis* 15, 124-139.
- Roodman, D. (2009): How to do xtabond2: An introduction to difference and system GMM in Stata, *The Stata Journal* 9(1), 86-136.
- Rosenthal, S. and W. Strange (2004): Evidence on the Nature and Sources of Agglomeration Economies, *Handbook of Urban and Regional Economics* 4, Amsterdam: Elsevier.
- Saito, H. and M. Gopinath (2009): Plants' self-selection, agglomeration economies and regional productivity in Chile, *Journal of Economic Geography* 9, 539-558.
- Slaughter, M. (2001): International trade and labor-demand elasticities, *Journal of International Economics* 54(1), 27-56.
- Südekum J. (2006): concentration and specialization trends in Germany since reunification, *Regional Studies* 40(8), 861-873.
- Wagner, J. (2007): Exports and productivity: a survey of the evidence from firm-level data, *The World Economy* 30(1), 60-82.
- Windmeijer, F. (2005): A finite sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics* 126(1), 25-51.

Appendix

Table 1: Data sources for the regional variables

Variables	Data source
Population	Federal Statistical Office
Area in km ²	Federal Statistical Office
GDP	Statistical Offices
Accessibility	Federal Office for Building and Planning
Number of employees	Federal Employment Office
Number of plants and employees	Establishment History Panel of the IAB

Table 2: Overview over the variables

variables	abbreviation	description
<i>dependent variable</i>		
Output	<i>y</i>	Value added (in Mill Euro)
<i>Basic production function variables</i>		
Labor	<i>l</i>	Employees (in full-time equivalents) times working hours
Capital	<i>k</i>	Capital stock (in Mill Euro)
<i>Region-specific variables</i>		
Specialization	<i>spec</i>	Krugman specialization index (KSI)
Concentration	<i>conc</i>	Localization quotient (LQ)
Diversity	<i>div</i>	Hirshman-Herfindahl index across sectors
GDP density	<i>gdparea</i>	GDP per km ²
Accessibility	<i>access</i>	Driving time by car to nearest highway (in minutes)
<i>Plant-specific control variables</i>		
Export activities	<i>export</i>	Share of sales abroad on total sales in percent
Works council	<i>counc</i>	dummy: 1= works council
Dependency structure	<i>struct1</i>	dummy: 1= independent
	<i>struct2</i>	dummy: 1= branch
	<i>struct3</i>	dummy: 1= head office
Property structure	<i>fprop</i>	dummy: 1= foreign property
Age	<i>young</i>	Dummy: 1= plant founded after 2001
Sector	<i>WZ</i>	Dummies for WZ=15 to 74
East-West	<i>east</i>	dummy: 1= Eastern Germany

Table 3: Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
y	7.928	42.352	1,05	1.524.292
ln y	7,27	1,72	0,05	14,24
k	36.900.000	314.000.000	232,14	16.400.000.000
ln k	14,39	2,24	5,45	23,52
spec	0,62	0,11	0,37	1,32
ln spec	-0,49	0,17	-1,00	0,27
conc	2,95	8,48	0,00	202,03
ln conc	0,30	1,07	-6,22	5,31
div	0,03	0,02	0,02	0,38
ln div	-3,59	0,28	-4,05	-0,97
gdparea	25,07	37,52	0,53	236,19
ln gdparea	2,14	1,52	-0,63	5,46
access	13,77	9,49	0,40	63,67
export	11,58	22,27	0	100
counc	0,35	0,48	0	1
struct2	0,15	0,36	0	1
struct3	0,11	0,31	0	1
fprop	0,08	0,26	0	1
young	0,05	0,22	0	1