

**FIRM STRUCTURE AND THE LOCATION
DECISION OF GERMAN
MANUFACTURING FIRMS –
EVIDENCE FROM OFFICIAL FIRM-
LEVEL DATA**

Astrid Krenz

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Firm structure and the location decision of German manufacturing firms – Evidence from official firm-level data

Astrid Krenz, University of Göttingen*

Abstract

This paper uses a comprehensive, official firm-level dataset for German manufacturing firms to investigate the location decision of new firm activity in the German regional economy, differentiated by firm structure. The rich regional dimension of this dataset is investigated for the first time in regard to the location choices of firms. Results reveal that agglomeration economies play a significant role for small firms, but not for medium-sized and large firms. Whereas the market potential exerts a significant positive impact for all firms, labor costs do not exert a significant impact on large firms' location decisions.

Keywords: Firm location, Regional economy, Agglomeration economies

JEL-Code: R11, R12

Acknowledgements and Notes. The author has programmed several Stata codes for the analysis. The data access has been acquired in the form of controlled remote data processing. The firm-level data are strictly confidential, access can be attained by following the guidelines at <http://www.forschungsdatenzentrum.de/datenzugang.asp> . The author is grateful to Florian Köhler from the Statistical Office in Hannover who was very helpful in managing the dataset and running the Stata codes on the confidential dataset at the Statistical Office in Hannover. The author is also grateful to Rita Hoffmeister from the Statistical Office in Hannover and to Michael Rößner from the Statistical Office in Halle-Saale. Financial funding from the Lindemann foundation for this project is gratefully acknowledged.

*Department of Economics, University of Göttingen, Germany, Platz der Göttinger Sieben 3, 37073 Göttingen, astrid.krenz@wiwi.uni-goettingen.de

Introduction

More than 25 years after German reunification a divide in terms of firms' economic performance between the East and West of Germany still exists. A greater share of manufacturing firms in the West than in the East exports, human capital intensity as measured by wages per employee is higher in the West, as well as the size of firms measured in terms of the number of employees (Wagner 2008). After the Fall of the Wall, firms in the East had to be restructured and privatized, many of them were closed down.

This asymmetrical development of firm activities gives rise to a variety of research questions: What are the characteristics of those firms that operate either in the West or East of Germany, in the metropolitan areas or in the periphery? Which factors can explain the location decision of firms across the Eastern and Western German regions? Which role do agglomeration externalities (benefits accruing from co-location of firms due to specialized intermediate inputs supply, labor market pooling and knowledge spillovers) play in that process?

A formalization of the causes and determinants of localization is provided in the New Economic Geography, a research field which was initiated by Paul Krugman (1991). In the models, agglomeration of economic activity is explained through an interplay of increasing returns to scale and transport costs. Supplier and demand linkages yield cost reductions and better market access and enforce the agglomeration of economic activity. In addition, further factors have been revealed by the prior literature to play an important role for the agglomeration of economic activity, for example the infrastructure or institutional factors. Moreover, advantages from producing close to investors from the home country have been revealed (Figueiredo et al. 2002; Crozet et al. 2004).

A central focus of this paper lies on examining the factors that determine the location choice of first time investments of firms in the German economy. It enriches the research literature by running a micro-level analysis on a comprehensive firm-level dataset from the German Federal Statistical Office and makes use of its detailed regional information for the first time.

The paper is organized as follows. The next part describes the related literature. Part three deals with the methodological approach. Part four explains the data operationalization and variables. Part five covers the empirical analyses and results. The last part concludes.

Related Literature

The literature on location decisions can be grouped into studies that investigate the localization in the home country by domestic firms (see among others Figueiredo et al. 2002; Carlton 1983) and into studies that focus on the location decisions of foreign owned firms and FDI in host and home countries (Crozet et al. 2004; Head et al. 1995; Pusterla and Resmini 2007; Procher 2011; Disdier and Mayer 2004, e.g.).

In terms of domestic localization, the study by Figueiredo et al. (2002) investigates the investment decisions of Portuguese entrepreneurs. The authors find that the investor's prior base of economic activity is an important factor. A homefield advantage exists, making the investors accept higher labor costs to produce there. Agglomeration economies play an important role in compensating for a non-homefield localization disadvantage. Carlton (1983) investigated the location decision of firms in the US and found evidence of the importance of agglomeration economies.

As for the literature on inward foreign investments, in a study on FDI in France, Crozet et al. (2004) show that investors locate in regions close to their home market and that agglomeration effects are important. Head et al. (1995) find that Japanese investors tend to co-locate with previous Japanese investors in the US. Pusterla and Resmini (2007) investigated the location choices of foreign firms in selected CEE countries, namely Bulgaria, Hungary, Poland and Romania. They find that agglomeration economies and market demand are important factors as well as low labor costs.

In terms of outward foreign investments, Vivien Procher finds in her analysis on the location decision of French first-time investments across Europe, North America and North Africa that investments of firms are fostered by cultural proximity to France, that manufacturing firms prefer to locate in the new Eastern European countries and that agglomeration effects are important (Procher 2011). Disdier and Mayer (2004) investigated location choices of French multinational enterprises across Western and Eastern European countries. They find that market size, agglomeration economies and institutional factors are important for the location decision.

Only a few studies address both firm-level and regional-level characteristics for the location choice (Arauzo-Carod and Manjon-Antolin 2004; Procher 2011; Foreman-Peck and Nicholls 2015, e.g.). This might be due to issues about data availability and comprehensiveness. Arauzo-Carod and Manjon-Antolin find that the size of a firm matters, as in large firms objective decision-making is taking place, while in smaller firms the decision-

making process is led by preferences of the entrepreneur. In the following analysis, the location decision will also be investigated according to the size of firms.

Among the evidence that has been found for the German economy, enterprise data from the German Bundesbank have been used to investigate the localization of German multinational firms abroad, finding a significant impact of agglomeration effects (Buch et al. 2005), and a focus on special sectors, for example the machine tool industry and relocation decisions after World War II (Buenstorf and Guenther 2010) or high technology sectors in West-Germany (Bade and Nerlinger 2000) has been established. Using official firm-level data from the German Federal Statistical Office instead, bears several advantages, as it allows for a comprehensive analysis of all firms' activities as firms are obliged to report, so the analysis is possible to a much wider extent than with other firm-level datasets. Moreover, a further benefit is the dataset's rich regional information.

Methodology

The aim of the analysis is to investigate the location choice of a firm and the determinants for the location decision. The investor chooses among several regions where he can set up his firm activity. A discrete choice model will apply to this decision problem.

The conditional logit approach from McFadden (1974), which is frequently used in the literature and has been introduced to the location literature by Carlton (1983), will be applied for the analysis. The theoretical foundation is based on a profit maximization problem of the firm. An investor i will choose a location r over a location s , if that location's expected profit is higher than in region s :

$$\pi_{ir} > \pi_{is}, \quad \forall s, s \neq r, \text{ and } \pi_{is} = \gamma_{is} + \varepsilon_{is}.$$

In this framework, an investor's profit π consists of a systematic and a stochastic part. The systematic part γ is a deterministic function of observable characteristics exerting influence on profits. It can be specified as a linear combination of region-specific attributes. A set of coefficients – for a set of explanatory $l = 1, \dots, L$ region-specific variables X – can be estimated:

$$\gamma_{is} = \sum_{l=1}^L \beta_l X_{ls}.$$

The stochastic part ε captures non-observed heterogeneity and random components that drive the investor and his investment decision.

The dependent variable in the discrete choice model is a binary variable that carries the value one if a region is chosen for a firm's investment and zero if otherwise. The probability that a firm i chooses a location r is given by:

$$p_{ir} = \frac{e^{x_r' \beta}}{\sum_{m=1}^M e^{x_m' \beta}} \quad , m = 1, \dots, r, \dots, s, \dots, M \quad \text{and } i = 1, \dots, N .$$

The conditional logit model makes a strong assumption about inter-regional independence: the probability ratio of two locations is said to be independent of any other third location (Independence of Irrelevant Alternatives). An alternative method taken in the literature is the nested logit model. The problem with the nested logit model, however, is that in the case that the number of observations in some nests is too small, the estimates will not be robust. This is the case for the German firm-data in its regional levels. The literature states that the conditional logit estimates will provide a good estimation in case one is interested in the average preferences rather than in predictions on the odds ratios due to varying regional attributes (Train 2004; Procher 2011).

Data and Variables

For the study official firm-level data from the German Federal Statistical Office and the statistical offices of the *Länder* (federal states) are taken. These data provide information about the enterprises in the German economy. Firms are obliged to report. A special feature of the data is that it contains detailed regional information about the location of the firm up to the community level (LAU-2). This makes the dataset unique, other firm-level datasets frequently used in the literature do not contain such a rich regional classification. The dataset covers a wide range of information on employment, revenues, investment, taxes, R&D expenditures and further costs. The regional dimension of the firm-level dataset offers potential for many analyses and new research outcomes, as it has not been used for comprehensive regional studies before. Official German Federal Statistical Office's firm-level data have been used so far to investigate micro-level issues for the relationship between exports and productivity, for example, without exploiting the detailed regional information (Wagner 2002, 2007).

The firm-level data that were used for this study is called the AFiD-Panel *Industrieunternehmen*. This panel consists of all firms in the sectors manufacturing, mining and quarrying and reports those firms which employ at least 20 employees. For the analysis, only the firms in the manufacturing sector were extracted. The firm-level dataset requires a special handling of the regional dimension. As in the German economy regional reclassifications occur frequently over the years - because communities are merged or split up -, this had to be coded by the author in the Stata program. For reasons of IT capacity limitations, the data were aggregated up to the NUTS3-level of administrative districts (*Landkreise* and *Kreise und kreisfreie Städte*).

In order to capture new firm activities, similar to Procher (2011) the argument of Roberts and Tybout (1997) is operationalized. Roberts and Tybout find that exporting firms who are out of the international market for more than two years have to bear similar re-entry costs as new entrants to the market. Thus, even if a firm was active in a year before 2011 it can be considered as a new entity, as it has to face high re-entry costs. This type of operationalization is necessary, as the German official firm-level data do not provide firm-demographic variables for scientific analyses, yet. In this paper, a new firm activity is defined as a firm operating in 2013 but not in 2012 or 2011.

The analysis will capture the location decisions of domestic firms across the German regional economy. The impact of localization determinants is examined for different subsamples of firms, grouped by firm structure into small, medium-sized and large firms. Small firms are defined as having less than 50 employees and total revenues of 10 mio euros and less. Large firms are defined as having 250 employees and more and total revenues of more than 50 mio euros. Medium-sized firm have between 50 and less than 250 employees and total revenues between 10 mio and 50 mio euros. For the final samples location choices by all, only the small-sized, medium-sized and large firms across 429 NUTS3-level regions are included.

The data on regional characteristics are on the one hand taken from the INKAR database of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). Further variables were taken from the Regional Database of the Federal Statistical Office. The data were merged to the firm-level dataset. The reference year for all the explanatory factors is 2012. It follows the idea that an investor will form his location decision based on the past year's regional attributes and not on the current year's ones. Moreover, this method will be less prone to endogeneity issues, which is also a common procedure undertaken in the literature. The regional level of the analysis is the NUTS3 level. Descriptive statistics of the variables can be found in Table A.1 and Table A.2 in the Appendix. A further description of the variables can be found in Table 1.

Table 1 List of Variables

Variable	Description	Data source
Dependent variable	Binary variable, equals 1 if the region was chosen as location, and 0 for the other regional alternatives, varying across firm and region	AFiD-Panel <i>Industrieunternehmen</i> , German Federal Statistical Office and the offices of the

		<i>Länder</i>
ln(population)	Total population in a given year (in logs)	Regional Database, German Federal Statistical Office
ln(GDP per employee)	GDP per employee, in 1000 euros, measured at market prices (in logs)	INKAR, BBSR
GDP growth	GDP growth rate, change from 2008-2013, in %, measured at market prices	INKAR, BBSR
agglomeration economies	Index out of the statistics on the number of firms, share of a region's number of firms in manufacturing given the number of firms over all regions in manufacturing divided by share of the number of all firms in that region given the number of all firms in Germany	Regional Database, German Federal Statistical Office, author's computations
	$\frac{f_{rM}}{(\frac{f_M}{f_r})}$ $\frac{f_{rM}}{F}$	
labor costs	Gross wages and salaries in euros per employee in the sectors Manufacturing, quarrying and mining (with at least 20 employees), measured in the month June	INKAR, BBSR
corporate tax	Corporate tax in euros per inhabitant	INKAR, BBSR
distance to agglomeration center by car	Average of the time (in minutes) to arrive at the 3 nearest (out of 36) agglomeration centers by car	INKAR, BBSR
East dummy	Dummy variable which equals 1 if the location is in Eastern Germany	author's computations

The following regressions will on the one hand control for factors that proxy for the market potential. One of the most important location determinants found in the literature is the market demand. The size of the population is one prominent way to capture market demand. A larger population in a region might result in a higher demand for firms' products. GDP per employee is another important variable proxying for the market potential. It is capturing consumers'

purchasing power and welfare and the functioning of institutions. The growth of GDP is further considered to capture the dynamics of economic progress and recessions.

In various studies, agglomeration economies have been found to exert an important influence on the location choice of firms (Carlton 1983; Figueiredo et al. 2002; Procher 2011). For the analysis, an index measure is computed from the statistics on the number of firms from the Regional Database of the Federal Statistical Office. It is formed as the share of the number of firms per region given the number of manufacturing firms over all regions divided by the number of all firms for the region given the number of all firms in Germany. The measure shows how much firm activity is present in a given region. Thus, it can be investigated if firms benefit from the clustering with other firms.

Labor costs are controlled for to capture cost factors for the location decisions of firms. The literature found ambiguous evidence about the effects of this variable. On the one hand, the cost argument would indicate a negative relationship with the location decision of a firm. Higher labor costs will reduce firms' profits and therefore firms will be less attracted to a region (see e.g. Pusterla and Resmini 2007, Procher 2011). On the other hand, higher wages might indicate a region's higher share of highly skilled workers, which would then indicate a positive relationship between labor costs and the location decision.

Another cost factor for the firm is corporate taxes. A negative relationship with the location decision of firms might be expected. However, several studies in the literature found a positive effect or no significant effect at all (Carlton 1983, e.g.). An explanation for that evidence might be that with a higher tax income, more public investment on information, communication and infrastructure can be undertaken (Bellak et al. 2009a, 2009b).

Further explanatory factors for the location choice entail the distance to centers of economic activity as measured by the average travel time in minutes by car to reach the three nearest agglomeration centers, and an East-West dummy variable.

Estimation Results

The estimation results from the conditional logit model for different firm samples investigating the determinants of the location choices of firms are displayed in Table 2. The baseline results are given first, the second set of results contains the variable capturing agglomeration economies.

Market demand, as measured in terms of the size of the regional population, has a positive impact on the location decision of all firms. This result is in accordance with previous evidence in the literature (Procher 2011, Crozet et al. 2004). The impact from GDP per

Table 2 Estimation Outputs

	All firms	All firms	Small firms	Small firms	Medium-sized firms	Medium-sized firms	Large firms	Large firms
ln(population)	0.6888**	0.3118**	0.7257**	0.2964**	0.683**	0.4232*	0.4937**	0.3702*
ln(GDP per capita)	0.8199**	0.7951**	0.9906**	0.9794**	0.4571	0.4326	0.1868	0.1595
GDP growth	-0.0124**	-0.0161**	-0.0211**	-0.0256**	0.0176	0.0154	-0.0135	-0.0144
agglomeration economies		0.0003**		0.0004**		0.0002		0.0001
labor costs	-0.0004**	-0.0003**	-0.0004**	-0.0003**	-0.0005**	-0.0005**	-0.00009	-0.00008
corporate tax	0.0002	0.0002	-0.0001	-0.0001	0.0004	0.0004	0.0009**	0.0009**
distance to agglo center by car	-0.0009	-0.0018	-0.0015	-0.0026	-0.0051	-0.0058	0.0052	0.0049
East dummy	0.096	0.0667	0.169*	0.1331	0.0516	0.0331	0.2616	0.2579
Observations	500588	500588	304854	304854	53196	53196	61039	61039
Wald Chi2	376.83	685.69	279.39	537.94	34.87	54.39	54.37	59.15
Prob>Chi2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Log pseudolikelihood	-8397.0697	-8380.135	-5095.214	-5082.069	-890.8965	-890.0596	-1023.457	-1023.211

Source: Data from German Federal Statistical Office and the statistical offices of the *Länder*, from INKAR/ BBSR, and from the German Regional Database, author's computations.

Notes: Reported are results from a conditional logit estimation. Dependent variable is the chosen NUTS3-region (1) and the alternative regions (0).

The number of observations corresponds to the number of firms x the number of regions for localization. * denotes $p < 0.10$, ** denotes $p < 0.05$.

employee is important for small firms only and GDP growth appears not to be a positive force for the location decision, at all.

Agglomeration economies exert a significant and positive impact for the small firms' location choices. Its impact for the medium-sized and large firms is not significant. Apparently, the smaller firms benefit from the externalities arising from the agglomeration of economic activity, whereas for large firms other factors are more compelling for a positive location decision.

Labor costs exert a significant impact on small and medium-sized firms' location decisions, but do not so for the large firms. The corporate tax is not having a negative influence, for the small and medium-sized firms not even a significant effect for the location choice. For the large firms, a positive effect exists. This might be explained by previous evidence found in the literature (see e.g. Carlton 1983), and point to a greater degree of transfers and investment into information, communication and infrastructure in a region, which might be considered highly relevant for large firms.

The distance to the agglomeration center by car is not significant for any firm sample. The results for choosing a location in the East of Germany are not robust. Adding the measure of agglomeration economies into the regression framework, the positive effect from the baseline specification for the sample of small firms disappears.

Conclusion

In this paper, the location decisions of new firm activities in the German regional economy, differentiated by the size of firms, have been investigated. The analyses revealed that higher market demand increases the probability of a location to be chosen for all firms, whereas higher labor costs are detrimental for the location decision of small and medium-sized firms. Small firm investors are attracted by the agglomeration of economic activity / of firms in a region.

The analysis revealed no clear picture whether firms prefer to locate rather in the East or West of Germany. For that reason, in a follow-up study the investment decisions between Eastern German and Western German firms are analyzed. Another direction of further research is to make use of the rich information at the firm-level, to control for interaction effects between firm-level and regional-level attributes to gain a deeper understanding of driving forces of firms' investment and localization decisions.

The historical German division into East and West was found in the prior literature to be still relevant, as it can be seen by different performance indicators that are still diverging

between the East and the West, like exporting activity, wages per employee or the size of firms. A process of restructuring of firms was set in place after German reunification, which creates a need to further investigate to what extent regional industry clusters developed, how agglomeration externalities across firms emerged and how location decisions of firms across the German regional economy are determined. The regional localization of firm activity is important for regional economic growth and welfare and thus for the regional development of one of the most vital economies in the European Union.

References

- Arauzo-Carod, J.; Manjon-Antolin, M. (2004), Firm Size and Geographical Aggregation: An Empirical Appraisal in Industrial Location, *Small Business Studies*, 22, 299-312.
- Bade, F.; Nerlinger, E. (2000), The spatial distribution of new technology-based firms: Empirical results for West-Germany, *Papers in Regional Science*, 79, 155-176.
- Bellak, C.; Leibrecht, M.; Damijan, J. (2009a), Infrastructure endowment and corporate income taxes as determinants of foreign direct investment in Central and Eastern European countries, *World Economy*, 32, 267-290.
- Bellak, C.; Leibrecht, M. (2009b), Do low corporate income tax rates attract FDI? Evidence from Central and East European countries, *Applied Economics*, 41, 2691-2703.
- Buch, C.; Kleinert, J.; Lipponer, A.; Toubal, F. (2005), Determinants and effects of foreign direct investment: Evidence from German firm-level data, *Economic Policy*, 41, 51-110.
- Buenstorf, G.; Guenther, C. (2010), No place like home? Relocation, capabilities, and firm survival in the German machine tool industry after World War II, *Industrial and Corporate Change*, 20, 1-28.
- Carlton, D. (1983), The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables, *Review of Economics and Statistics*, 65, 440-449.
- Crozet, M.; Mayer, T.; Mucchielli, J. (2004), How do firms agglomerate? A study of FDI in France, *Regional Science and Urban Economics*, 31, 27-54.
- Disdier, A.; Mayer, T. (2004), How different is Eastern Europe? Structure and determinants of location choices by French firms in Eastern and Western Europe, *Journal of Comparative Economics*, 32, 280-296.
- Figueiredo, O.; Guimaraes, P.; Woodward, D. (2002), Home-field advantage: location decisions of Portuguese entrepreneurs, *Journal of Urban Economics*, 52, 341-361.

- Foreman-Peck, J.; Nicholls, T. (2015), Inter-regional mobility of entrepreneurial SMEs, *Annals of Regional Science*, 54, 57-87.
- Head, K.; Ries, J.; Swenson, D. (1995), Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States, *Journal of International Economics*, 38, 223-247.
- Krugman, P. (1991), Increasing returns and economic geography, *Journal of Political Economy*, 99, 483-499.
- McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior, in Paul Zarembka (ed.), *Frontiers in Econometrics*, New York: Academic Press, 105-142.
- Procher, V. (2011), Agglomeration effects and the location of FDI: evidence from French first-time movers, *Annals of Regional Science*, 46, 295-312.
- Pusterla, F.; Resmini, L. (2007), Where do foreign firms locate in transition countries? An empirical investigation, *Annals of Regional Science*, 41, 835-856.
- Roberts, M.; Tybout, J. (1997), The decision to export in Colombia: an empirical model of entry with sunk costs, *American Economic Review*, 87, 545-556.
- Train, K. (2003), *Discrete choice method with simulation*, Cambridge University Press, Cambridge.
- Wagner, J. (2008). A note on why more West than East German firms export, *International Economics and Economic Policy*, 5, 363-370.
- Wagner, J. (2007). Exports and Productivity: A Survey of the Evidence from Firm Level Data, *The World Economy*, 30, 60-82.
- Wagner, J. (2002). The Causal Effects of Exports on Firm Size and Labor Productivity: First Evidence from a Matching Approach, *Economic Letters*, 77, 287-292.

Appendix

Table A1 Descriptive Statistics – regional characteristics

	Mean	St Dev	p1	p25	p50	p75	p99
ln(population)	12.042	0.6481	10.5887	11.6436	12.0058	12.4376	13.9223
ln(GDP per capita)	4.0888	0.1504	3.7887	4.0037	4.0877	4.1573	4.5992
GDP growth	10.5425	7.1594	-6.5	6.1	10.3	14.8	29.5
labor costs	3316.219	720.5841	2037.6	2803.2	3273.1	3692.5	5446.3
corporate tax	458.1197	277.463	132.8	292.2	385.1	527.1	1828.6

distance to agglomeration center by car	97.1012	25.5642	48	80	98	113	157
East dummy	0.217	0.4122	0	0	0	0	1
agglomeration economies	708.2821	543.8311	124	380	552	894	2458

Notes: p1, p25, p50, p75 and p99 refer to the 1st, 25th, 50th, 75th and 99th percentile of the distribution of the regional characteristic.

Table A2 Descriptive Statistics – firm-level characteristics

	Mean	St Dev
number of employees	92.7709	53.6647
Total revenues	2.31e+07	1.48e+08

Notes: The minima and maxima values cannot be displayed as it violates rules of confidentiality.