What Drives the Location Choice of New Manufacturing Plants in Germany?*

Astrid Krenz†

Abstract. More than 30 years after German reunification, a persistent gap in different firm performance measures exists between East and West Germany. In this paper I focus on the differences in new German manufacturing plants' location choices across the German district-free cities and districts and investigate its regional determinants. For that purpose, I construct a novel, rich regional- and plant-level dataset based on the Official Firm Statistics from the German Federal Statistical Office and the Offices of the Laender. The analysis provides first-time evidence regarding how in particular the location decision of plants in the German economy is influenced by regional road infrastructure as well as regional structural funding. The effects are economically important and significant. The results reveal that a 10 percent increase in plant agglomeration increases the odds of a new plant to locate in the region by 12 percent. A 10 percent decrease of travel time on roads increases the odds of a plant to locate by 4 percent in Germany overall, by 7.6 percent among East German regions and by 26.5 percent in particular for large plants in the East German regions. A 10 percent larger population increases the odds to locate by 8.7 percent. A 10 percent increase in regional structural funding for infrastructure purposes increases the odds to locate in a region in East Germany by 8.3 percent in particular for large plants. Policy implications emerge that address in particular the improvement of infrastructure and support to reap the benefits that arise from agglomeration externalities.

Keywords: Firm location choice, regional road infrastructure, Germany, agglomeration economies, regional structural funding, East-West gap, conditional logit, nested logit.

JEL: D22, L25, R11, R12.

---

* I am thankful for support by an EU Marie Curie Cofund / Durham International Junior Research Fellowship under EU grant agreement number 609412. I am further thankful for comments from (blinded for review). The Stata Code that was programmed for the analyses is available upon request. The Stata Codes that I programmed were run on the computers in the Statistical Office in Hannover by Florian Koehler. Florian Koehler and Rita Skorka were in charge of the check of validity of the original data, the check of validity of my constructed dataset and the checks of confidentiality regulations of my results. The data used are confidential but not exclusive and come at a cost, further information can be found at www.forschungsdatenzentrum.de/nutzungsbedingungen.asp. Remote data access applied.

† University of Sussex, Digit Research Centre, Jubilee Building, Falmer, BN1 9SN, Brighton, UK; email: a.m.krenz@sussex.ac.uk.

1. Introduction

One of the most important questions for politics is where new firms set up their economic activity and locate. Firms are local providers of employment. Regions that have successful firms operating can offer jobs and generate higher GDP. The question of where firms locate has been prominent in German politics for many years and has been given renewed interest with the recent plan for a "German Industrial Strategy 2030".[1] A longer-lasting debate exists about underinvestments in the German economy, involving spending on infrastructure (i.e. a lack of investment in roads, digital as well as energy infrastructure), support for innovations and research, as well as support for new firms in terms of easing regulations, bureaucracy, and access to finance ("Expertenkommission Staerkung von Investitionen in Deutschland", 2015). These issues bear an important influence on investment decisions and setting-up of firm activities. An evaluation of the impact that infrastructure and subsidies have on the location choice of plants may enrich the discussion about which investments are required to prepare the German economy and society for future developments in a changing environment of technologies, climate, international trade, and political relations.

The focus of this paper is to examine the factors that determine the location choice of new manufacturing plants in the German economy. For that purpose, I constructed a novel, comprehensive, and rich regional- and plant-level dataset that is based on official firm data from the German Federal Statistical Office and the Offices of the Laender. This paper links for the first time detailed regional information from external sources with the comprehensive official firm data to investigate German plant location decisions. Given the richness of the data, I will be able to provide, for the first time, evidence on how the regional road infrastructure and regional structural funding impact the location choice of plants between German regions.

The German case is special in so far as through almost 40 years of division between East and West after World War II, firm performance between East and West Germany became different.[2]

[1]The strategy paper discusses how large enterprises, the so-called "champions", can be supported to flourish (BMWI, 2019). The German government aims to cope with the rising economic power from the Asian countries, especially China, and the USA. It envisions that only a strong Europe - and Germany - can master the competition arising from the USA and Asia. A major goal of the strategy would be that the government intervenes when foreign ownership is taken over in relevant industries. Given that this will limit free competition in a market economy, this led to a controversial debate and criticism among firm corporations, politicians as well as scientists.
[2]Investigating the German case is interesting, and given the high quality of the firm level dataset a feasible undertaking. Other divides are also an important issue in other European countries, think for example about the North-South divide in Italy, in the UK, in Spain or between capital and peripherial regions such as countries like the Czech Republic or Romania.

A divide in terms of firms' economic performance between East and West Germany still exists more than 30 years after German reunification. With the 30th anniversary of the fall of the Wall, a significant amount of attention in research as well as in the media has been drawn to regional disparities that appear to remain persistent across the German economy. Still, a higher share of manufacturing firms in the West than in the East exports, wages per employee are higher in the West than in the East, as is the size of firms measured in terms of the number of employees. Further evidence shows that productivity measures still differ across East and West Germany (Burda and Severgnini, 2018) and a lower industry output is produced by East German than by West German firms. The differences in firm performance appear to be persistent. As of today, it is not fully understood why the gap is not closing over time. Previous literature has found outward migration of highly skilled workers from the East German regions as well as low-quality infrastructure to affect lower East German firm productivity (Burda and Hunt, 2001). With the present contribution, I provide for the first time a comprehensive analysis of why plants prefer to choose a location either in West or East German regions, more precisely which regional factors affect the location choice.

A look back into Germany's history helps to reveal where we stand today. A crucial element of a market economy is the investor's free decision of where to invest and to set up production facilities. In the former Democratic Republic of Germany, which was a system of planned economy and state socialism, the location decision of a firm was undertaken centrally by the East German government. The enforcement of such localization can be detrimental to productivity because it hinders investment flowing to its most efficient use. An example for forced firm localization is the setting up of an iron and steel production facility in Eisenhuettenstadt (close to the Polish border) rather than in Rostock (close to the Baltic Sea coast) (Gayko, 2000); the Soviet Union did not approve the setting up in Rostock. It wanted to secure the provision with Soviet ores. Establishing the facility on the Baltic Sea coast would have made Scandinavian ore provision less costly and more easily available by transport routes, an outcome that the Soviet Union wanted to prevent.

In the time between World War II and German reunification, firms in East Germany were mainly state- and public-owned (see e.g. Sleifer, 2006). After the fall of the Wall, a process of liberalization and privatization took place. Many firms in East Germany were shut down or downsized. Firms were acquired by West German competitors or new plants from West

German or other external investors were established in the East German regions. Many firms were newly founded, with a peak in 1991. In the following years, though, investments slowed and a productivity gap between East and West Germany persisted. Administrative and strategic as well as research and development departments were underrepresented in East Germany. Whereas large firms dominated East Germany before the fall of the Wall, it was mainly small fims that were founded after the fall (Sleifer, 2006). The transformation process not only affected the structure of firms, but it also had immense impact on employment and institutions and invoked a structural change in the East German regions. Many employees were laid off. Immediately after the fall of the Wall, state-subsidized programs for short-term work and early retirement were set up, but they stopped some years later. Public administration had to be restructured, and production structures in East Germany had to be changed and follow market economy principles.

The asymmetrical development between West and East Germany gives rise to a variety of research questions: Which factors can explain the location decision of plants in Germany? Do these determinants differ across East and West German regions? What role do agglomeration externalities, the regional infrastructure and regional funding play in that process? Which plants operate either in the West or in the East of Germany, and what are their characteristics?

This paper presents new evidence for the location decisions of new manufacturing plants and disentangles the driving factors of plant location choices in East and West German regions. To the best of my knowledge, it appears to be the first study that investigates the issue using the official firm data from the German Federal Statistical Office and the Offices of the Laender. I investigate a wide range of influential factors that matter for the location decision, such as agglomeration externalities, human capital, cost factors, regional road infrastructure, and regional structural funding. I matched the plant-level data with regional data from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) through its INKAR database (Indicators and Maps for Spatial and Urban Development) and the Regional Database of the German Federal Statistical Office (GENESIS). German law mandates that all enterprises and plants in the German economy have to report for the official statistics, and the register of firms captures the entire population of enterprises and plants in the German economy.[3] The comprehensiveness of the data, which also include the plant's location information, allows investigation of the decision to locate depending on regional characteristics. For that purpose, I exploit the variation of regional characteristics across NUTS III regions, the district-free cities and districts (*Landkreise and kreisfreie Staedte*) in Germany, to explain the location decision of a new manufacturing plant. To the best of my knowledge, this is the first time that the official German firm-level data are applied for investigating the location choices of firms. Other datasets for the German economy are basically survey-level data. Their number of observations is reaching at most 0.5 percent of the overall population of firms. However, the dataset that I use, basically covers all of the plants and enterprises in the German economy, i.e. about 3.5 millions. With this official firm-level dataset comes a wide range of further variables, which allows for a thorough investigation of the firm-level location choices. A conditional logit model is used for the estimations as well as a nested logit model for robustness checks. The conditional logit model controls for effects of the groups, which is here the characteristics that are given at the plant-level. Therefore, any plant-level specific fixed effects and industry-level fixed effects are controlled for by design of the econometric method, ruling out omitted variable bias. Furthermore, regional fixed effects are included into the regressions.

The main results indicate that an important factor for a plant's decision to choose a production location is regional infrastructure in terms of travel time on roads and consequently the accessibility of a region. How long it takes to reach agglomeration centers by car/ trucks matters for transporting goods. It is also important in terms of accessibility and local amenities for consumers. Moreover, agglomeration externalities as measured by agglomeration of plants per region, are found to bear a significant positive impact for the location decision of plants. The results further show that a larger amount of population per region is a major driving force for plants to locate. Furthermore, I can demonstrate for the first time that a higher amount of regional structural funding is important for the location choice of manufacturing plants across East German regions.

The paper is organized as follows. The next section reviews the previous literature. The third section describes the data and selection of variables. The fourth section explains the empirical analysis and discusses the results. The last section concludes.

[3] Excluded are only those from sectors A, O, T, and U according to ISIC rev. 4; these are the sectors of agriculture, forestry and fishing, public administration and defense, and other services.

2. Background Literature

The first documentable writings about regional location decisions of economic activity reach far back, to the year 1890, when Alfred Marshall discussed the externalities that firms will benefit from clustering together. These external effects imply cost savings for a firm and can be differentiated between localization and urbanization economies.[4] The benefits of localization comprise a local base of supplies to obtain cheap and adequate inputs for production, a local labor market to acquire qualified employees, and an ease of exchange of knowledge (Marshallian externalities). Urbanization economies imply benefits from a broader range of economic activity and market size (Jacobian externalities).[5] Early contributions of location decision studies that use firm data are Carlton (1983) and Bartik (1985).

A formalization of the causes and consequences of localization is provided by the New Economic Geography (Krugman, 1991). In these 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 shown to play a role in the agglomeration of economic activity; for example, institutional factors such as infrastructure as well as taxes and subsidies (Arauzo-Carod and Manjon-Antolin, 2010). Further evidence for these factors has been established by e.g., Glaeser and Kerr (2009) who find that new startups in the USA are set up according to input supplier and customer linkages and lower costs. Rosenthal and Strange (2003) found that agglomeration externalities are an important driving force for the birth of new establishments. Henderson (1994) established an econometric model to analyze industrial locations in Brazil, and finds evidence for industry subsidies to influence location patterns of industries and to increase the number of firms in cities.

The previous literature can broadly be divided between investigations of the location decisions of domestic firms in the home country (e.g. Carlton, 1983; Bartik, 1985; Arauzo-Carod and Manjon-Antolin, 2004; Hansen, 1987) and the location decision of foreign-owned firms and FDI at home and abroad (e.g. Guimaraes et al., 2000; Head et al., 1995; Devereux and Griffith, 1998; Procher, 2011; Buch et al., 2005; Becker et. al, 2005). Carlton (1983) investigates the

[4] The influence of externalities has been investigated among others by Glaeser et al. (1992) and Porter (1990).

[5] Jacobs (1969) elaborated the idea that diversity, industrial variety, and externalities derived from other industries rather within industries are conducive to growth. It is the differences between people, occupations, and industries that encourage knowledge and technological spillovers and therewith innovations and growth.

location decision of new plants in the USA. His results reveal that agglomeration economies as measured by the number of production hours are exerting a positive and significant impact on the location choice. He finds no significant impact of wages and taxes, though. Bartik (1985) investigated the location decisions of new manufacturing plants in the USA. Using a conditional logit model, he finds that unionization and an increase in a state's corporate income tax and the business property tax rate lead to a decline in the number of new plants. Arauzo-Carod and Manjon-Antolin (2004), using a dataset on new Catalan establishments, show that in large establishments, more objective decision-making regarding location decisions is taking place, whereas in smaller establishments, the decision-making is led to an increased extent by preferences of the entrepreneur. They find a positive and significant impact of urbanization economies as measured by the total number of workers and negative effects from diseconomies. Hansen (1987) investigated the location decision of manufacturing plants in Brazil. He found that agglomeration economies strongly impact the location decision. He found no evidence for wages, though, but an impact of distance that would point to transport costs to matter for the location decision.

Regarding foreign direct investments and foreign-owned firms, Guimaraes et al. (2000) find that agglomeration economies in terms of localization and urbanization economies are important factors that determine the location choice of new foreign-owned plants in Portugal. Head et al. (1995) find that for Japanese investors in the USA in particular localization economies bear an important impact for the location decision. Devereux and Griffith (1998) investigate the location decisions of US multinational firms. They report that agglomeration economies are an important factor for the location choice. Once the firms have made their decision to produce in the European market, the effective average tax rate is found to influence the location decision within the European countries. Procher (2011) investigates the location decision of French new investments across Europe, North America, and North Africa. She finds that establishments are driven by a higher firm agglomeration, a higher market demand and cultural proximity to France. Higher labor costs and a larger distance between the headquarter and the foreign location are detrimental for the location decision. Among the evidence that has been found for the German economy, enterprise data from the German Bundesbank have been used to investigate the investment decision of German multinational firms abroad (Becker et al., 2005;

Buch et al., 2005). Becker et al. find that German multinational firms tend to locate in high-skilled labor abundant countries - but they find no impact for Swedish multinationals - and a larger wage gap between Germany and Central and Eastern European countries is associated with fewer jobs in the parent and more jobs in the affiliate firm. Buch et al. investigate the sales of German firms' foreign affiliates and find a negative impact from distance, a positive impact for market access, and a significant impact of agglomeration economies as measured by the number of German firms in a given foreign market.

3. Data

For the analysis, I use official firm data that are maintained by the German Federal Statistical Office and the Offices of the Laender. I merged data on manufacturing plants (AFiD *Industriebetriebe*) with data from the register of firms (*Unternehmensregister*, abbr. URS), and external regional data from the BBSR through its INKAR database, as well as from GENESIS.[6] The official firm data are comprehensive and allow me to conduct detailed regional analyses. Because German law mandates that all firms have to report to the official statistics, the data cover the entire population of plants and enterprises in the register of firms, which differentiates this dataset from other German firm datasets that are based on surveys or just on FDI.

A special feature of the data is that they contain detailed regional information about the location of a plant at the community level.[7] [8] This detail makes the dataset unique - other firm datasets frequently used in the literature do not contain such a rich regional classification. The regional dimension of the firm dataset offers potential for many analyses and new research. The regional information can be aggregated up to a higher NUTS level. Information about the NUTS-I level, the 16 federal states (*Bundeslaender*) in the German economy, is also contained within the firm data.

The observation unit used for the analysis is a manufacturing plant in the German regional economy in the year 2013.[9] The AFiD data on manufacturing plants entails all plants that have

[6] Access to the German firm data is project-specific. The mentioned datasets and variables therein are the ones that were approved by the Statistical Office to be used for my project.

[7] This was formerly the NUTS-V and is now named the LAU2 level.

[8] There were 11012 communities in Germany in 2018.

[9] Data provision by the German Federal Statistical Office and the Offices of the Laender takes considerable time. At the time of my application for the data (end of 2015), the last available year of data was for 2013. Thereafter, it took more than two and a half years until the Statistical Offices provided some first access to the data to me. Following data access, additional time is taken by the Statistical Offices for running the programmes and checks for confidentiality rules regarding the outputs.

at least 20 employees. I have data on all the plants of any enterprise. For the analysis, only the plants from the manufacturing sector were extracted. From the AFiD data on manufacturing plants, I retrieved the information regarding whether the plant was active in 2013.[10] To cover only new plants, I extracted those plants that were active in 2013, but not in 2012 or 2011, based on information from the register of firms URS (which has information on all enterprises and plants of all employment sizes). This procedure is based on arguments from Roberts and Tybout (1997) and has also been applied by Procher (2011). Roberts and Tybout found that exporting plants that are out of the market for more than two years have to bear similar (fixed) re-entry costs as new plants entering the market. Thus, for my analysis, even if a plant was active in a year before 2011, and not in 2012 and 2011, it can be considered as a new entity because it has to face high re-entry costs. This type of operationalization is necessary due to the fact that the official German firm statistics do not provide firm-demographic variables for scientific analyses, yet. My final sample comprises 1721 new manufacturing plants in 2013, having at least 20 employees.

The firm data require special handling of the regional information. As in the German economy, regional re-classifications frequently occur over the years - because communities are merged or split into new ones or new regional identifiers are allocated - this has to be taken care of. I used information on the reforms and changes (*Kreisreformen und Gebietsstandsaenderungen*) from the websites of the German Federal Statistical Office. For reasons of computability and IT capacity, I aggregated the data up to the NUTS-III level of the districts and district-free cities.

The analysis captures the location decisions of plants in the German regional economy. The impact of determinants is also examined for different subsamples of plants, namely small- and medium-sized plants as well as large plants. Small- and medium-sized plants are defined as having less than 250 employees and total revenues up to 50 million euros. Large plants are defined as having 250 or more employees and total revenues of more than 50 million euros.

I matched a range of regional variables to the firm data, on the one hand from the BBSR through its data portal INKAR and from GENESIS. The regional data come at the NUTS-III level. A total of 402 district-free cities and districts is considered in the regressions.

The explanatory factors are lagged by one year, such that the reference year for all the regional explanatory factors is 2012. It follows the idea that an investor will form his location decision

[10] To be considered as an active operating plant, it has to have at least one employee who is covered by social insurance.

based on the past's regional attributes and not on the current year's ones. Moreover, this method will be less prone to endogeneity issues, and it is also a common procedure undertaken in the regional economics literature. A list of variables, their data sources, and a description can be found in Table 1.

The dependent variable is binary and equals 1 for a region that was chosen as a location by a plant and 0 for the other regions that a plant did not choose to locate in. Hence, computability limitations are reached quickly.[11] For my analyses, more than 690000 observations are obtained, given that 402 regions times 1721 plant choices are investigated. The number of observations finally reduces when explanatory factors are not available for a region from the BBSR/ INKAR or the GENESIS database.

For the following regression analyses, I used a set of regional explanatory variables that were found to be important location drivers in the previous literature (see e.g. Procher, 2011, and see Table 1 for a summary). In various studies, agglomeration economies have been found to exert an important influence on the location choice of firms (e.g. Carlton, 1983; Guimaraes et al., 2000; Procher, 2011). For the analysis, I follow the literature and used a measure of the number of plants per population in a NUTS-III region to proxy for agglomeration economies. This measure shows how much firm activity is present in a given region and enables investigation of whether plants benefit from the presence of other plants. Further, the size of the region's population is controlled for. It is, on the one hand, an indicator of local market demand because a larger population in a region demands more products from firms. On the other hand, population has been used in the prior literature either to capture overall agglomeration economies or in particular urbanization economies that result from the regional presence of other facilities and industries and which measures economic diversity. Including the agglomeration of plants and the size of population thus proxies on the one hand for localization economies, the effects accruing from Marshallian externalities (supplier access, labor market pooling, and knowledge spillovers), and urbanization economies, the effects accruing from a diversification of many sectors, known as Jacobian externalities, on the other hand. GDP per capita and GDP growth are used to proxy for the region's welfare and to capture the dynamics of economic progress or decline. GDP growth is measured over the past year. On the cost side, labor costs in the manufacturing sector are controlled for. The literature provides ambiguous evidence for the effects of labor costs on

[11] Analyses are run on the statistical office's computer.

location decisions. 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. On the other hand, higher wages might indicate a region's higher share of highly skilled workers, a situation that would result in a positive relationship between labor costs and the location decision (Smith and Florida, 1994). Another cost factor for the firm is business tax (*Gewerbesteuer*).[12] A negative relationship with the location decision of a firm might be expected. However, several studies found a positive effect of taxes potentially indicating that more public investment is undertaken from the higher tax income or studies find no significant effect at all (see e.g. Carlton, 1983).

Among the institutional factors, the accessibility of a region measured by average travel time in minutes on roads to the three nearest agglomeration centers is considered. The road infrastructure variable is provided by the BBSR database and comes through its model for accessibility analyses (*Erreichbarkeitsmodell*) (see BBSR, 2019)[13]. In Germany, the largest share of the goods transfer (in ton km) is undertaken on the road, namely 70.0 percent in 2012, whereas 17.8 percent are undertaken on railways and 9.5 percent by shipping (see DIW Berlin, 2018).[14] For its analyses, the BBSR uses georeferenced information on a set of 662000 routes and 518000 knots for Germany and calculates travel times and km distances using ArcGIS. For road traffic, the BBSR uses information on type of streets (*Autobahn* (highway), *Landstrasse*, etc.), length of streets, velocity, and travel time (e.g. velocity profiles for both cars and trucks are used to model both passenger and goods transport on roads). The shortest travel times and/ or distances are computed and stored in time or distance matrices.[15] A negative impact for the location decision is expected from the road infrastructure variable because a higher value of this measure indicates lower accessibility of a region.

---

[12] In Germany, taxation of firms involves several items. A business tax that varies across communities has to be paid by firms. Further, firms either pay a corporate tax (*Koerperschaftssteuer*) of 15 percent or an income tax on profits, depending on the firm structure. Moreover, firms pay the so-called solidarity tax (*Solidaritaetszuschlag*) of 5.5 percent. From the INKAR database, I could retrieve information on the business tax and use it in the regression analyses.

[13] According to the BBSR definition, the variable captures the average travel time in minutes by passenger car to the closest 3 out of 36 agglomeration centers in Germany and the neighbouring countries.

[14] The transport rates did not change particularly towards 2017. 70.7 percent were transported on roads, 18.7 by railway, and 8.0 by shipping. In Germany in 2012, the length of public roads in 1000 km was 230.5 overall, 12.88 for *Autobahnen*, 39.6 for *Bundesstrassen*, 86.2 for *Landstrassen*, and 91.8 for *Kreisstrassen*. These figures did not change substantially until 2017 (latest available data) - the figures then were 229.9 for all roads, 13.01 for *Autobahnen*, 38.0 for *Bundestrassen*, 87.0 for *Landstrassen*, and 91.9 for *Kreisstrassen*.

[15] What their analyses show for example in terms of accessibility potentials of a region, is a peripheral accessibility for the very North-Eastern German regions, whereas the West German regions of the *Ruhrgebiet* display a high degree of centrality in terms of accessibility (BBSR, 2019).

Furthermore, a variable for regional structural funding measured by the long-term spending for regional infrastructure over the past ten years based on the Joint Task on Regional Economic Development by the German government (*Gemeinschaftsaufgabe zur Verbesserung der Regionalen Wirtschaftsstruktur*, abbr. GRW) plus co-funding from the European Regional Development Fund (ERDF) is considered. In order not to lose many observations within the regressions, I replaced a zero amount of funding in the dataset with a value of one euro. The GRW funding is a means to support investments in business and infrastructure for economic purposes in the German economy and to help redeem regional disparities (see e.g. IWH, 2018). The funding comes as an addition to basic funding that is undertaken by the firm itself. After German reunification, the East German regions mainly benefitted from the funding. The East German regions received approximately 20 billion euros in terms of GRW plus ERDF co-funding for economic-related infrastructure in the period between 1991-2012 (IWH, 2018). Therewith, 88 percent of the funding went to East German regions and 12 percent to West German regions. A positive impact is expected from this regional funding variable for the location decision of a manufacturing plant.

Other potential influential factors are captured by the inclusion of various fixed effects. The conditional logit method is known to control for effects at the level of the groups, which are in the context of this study the manufacturing plants. Therefore, any plant-level specific fixed effects and industry-level fixed effects are controlled for by design of the econometric method. Moreover, regional fixed effects to control for East and West German regions are included into the regressions.

Descriptive evidence is given in Tables 2, 3, and 4 and by Figure 1. One can find a highly differentiated picture of the distribution of plants across the German regions according to size. Figure 1 displays the distribution of plants across German district-free cities and districts. The variable that is given by the Regional Statistics INKAR (BBSR) is the number of plants per region. The graphics show both an East-West and a North-South divide in terms of firm activity. The highest numbers of plants per region (the darkest color displays the 25 percent regions with the largest numbers of plants) are found in West and in South Germany (the federal states North-Rhine Westfalia in the West and Baden-Wuertemberg in the South-West) rather than in East German regions. Moreover, firm activity is located in big city regions such as Berlin, the metropolitan areas of Munich, Hamburg, Cologne, the *Ruhrgebiet* regions, the South-Western regions around Stuttgart, Frankfurt, Hanover, and in the East German regions Dresden, Leipzig, and Jena. In East Germany, proportionally high numbers of smaller plants are operating.

Tables 2, 3, and 4 provide descriptive statistics for regional variables across the entirety of Germany, as well as East and West German regions. The results show that in East Germany, NUTS-III regions on average have a lower number of manufacturing plants (about 598 versus 732 in West Germany), a smaller population (209,893 versus 216,526 people), a lower GDP per capita (23 thousand versus 32 thousand euros per capita), a lower monthly manufacturing wage per employee (2559 versus 3516 euros), a lower business tax revenue per capita (278 versus 506 euros), a higher average travel time on roads to the three closest agglomeration centers (114 versus 93 minutes), and a higher amount of regional structural funding per capita (632 versus 23 euros).[16]

4. Empirical Analysis

4.1. **Methodology.** The aim of the analysis is to investigate the location choice of manufacturing plants in the German economy 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 applies to this decision problem.

On the one hand, the conditional logit approach from McFadden (1974), which is most frequently used in the location choice literature, is applied for the analysis. It allows modeling of a decision that has more than two discrete outcomes. For my analysis, the investor will make a choice for a location/ region (y=1) against a number of other regions that are not chosen as a location (y=0). 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:

$$\pi_{ir} > \pi_{is}, \forall s, s \neq r, \tag{1}$$

and

$$\pi_{is} = \gamma_{is} + \epsilon_{is}. \tag{2}$$

[16]The East German regions include the capital city Berlin.

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 the observable characteristics exerting influence on profits. It can be specified as a linear combination of region-specific attributes. A set of coefficients can be estimated for a set of explanatory variables X_l that vary across regions $l = 1, ..., L$:

$$\gamma_{is} = \sum_{l=1}^{L} \beta_l X_{ls}. \tag{3}$$

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 attains the value 1 if a region is chosen for a firm's location and 0 if otherwise. Under the assumption of interregional independence (IIA)[17], meaning that the probability ratio of two locations is independent of any other third location, McFadden (1974) shows that the probability that a firm i chooses a location r is given by:

$$P_{ir} = \frac{e^{X_r' \beta}}{\sum_{l=1}^{L} e^{X_l' \beta}} \tag{4}$$

with $l = 1, ...r, ..., s, ..., L$ and $i = 1, ..., N$.

The model is estimated by the maximum likelihood estimator. The interregional independence assumption is likely to be met the more aggregated the regional dimension considered is. The conditional logit model controls for effects of the groups, which is here the characteristics that are given at the plant-level. Therefore, any plant-level specific fixed effects and industry-level fixed effects are controlled for by design of the econometric method, ruling out omitted variable bias. Furthermore, regional fixed effects are included into the regressions.

For robustness checks, a nested logit model is also estimated. The idea behind nested logit is to generate groups of similar regions that an investor decides for his firm location, such that the interregional independence assumption is valid across nests. In other words, nests are formed that contain regions with similar attributes and that make regions substitutable for an investor,

[17] The IIA can be explained by the following example. When an investor has to decide to locate between two regions, the probability of a region to be chosen as a location would be 50 percent. When another region is available for location, the probabilities would have to be 33.3 percent for each location if the location choice was independent from the other region. However, it might be that two regions are rather substitutes for each other such that the probability to locate in region 1 remains at 50 percent, and the two other regions each have a probability to be chosen of 25 percent. The IIA would then be violated.

but the nests are not substitutes for each other. Nested logit modeling, however, is not a panacea because it is criticized for producing non-robust results: results vary depending on the chosen nest-structure.

Formally, the nested logit model can be described as follows (see Heiss (2002) and Cameron and Trivedi (2010)). An investor derives the following utility:

$$U_{kr} + \epsilon_{kr} = z_k^{'}\alpha + x_{kr}^{'}\beta_k + \epsilon_{kr} \tag{5}$$

Let τ_k be the so-called dissimilarity parameter, which measures the independence of alternatives in a nest k. It is that $\tau_k = \sqrt{1-\rho_k}$ with ρ_k the correlation of alternatives in nest k. The probability that a region (r, k) is chosen is then equal to the product of the probability of choosing nest k and the conditional probability of choosing r given k:

$$P_{kr} = P_k * P_{r|k} = \frac{e^{(z_k^{'}\alpha + \tau_k I_k)}}{\sum_{m=1}^{K} e^{(z_m^{'}\alpha + \tau_m I_m)}} * \frac{e^{(x_{kr}^{'}\beta_k/\tau_k)}}{\sum_{l=1}^{L} e^{(x_{kl}^{'}\beta_k/\tau_k)}} \tag{6}$$

In this formula, $I_k = ln(\sum_{l=1}^{L} e^{(x_{kl}^{'}\beta_k/\tau_k)})$ is the inclusive value which is the log sum of utilities generated from alternatives in nest k.

For my analyses, I attempted to find a grouping of regions that might be considered plausible for an investor's decision regarding where to locate in the German economy. The choice of nests (I chose 4 different groups of regions) for my analyses is displayed by Table 15. For the analyses, I assume that a new investor in a first step decides between either locating in: 1) one of the 25 largest German cities, 2) in one of the South-German district-free cities and districts from the federal states Bavaria and Baden-Wuertemberg, 3) in one of the regions in East Germany, or 4) in one of the regions in North-West Germany. In a second step, the choice for location of a district-free city or district is undertaken by the investor within the chosen nest. The variables in the matrix x are region-specific variables taken from BBSR/ INKAR, and the variables in z are case-specific and are the number of employees and turnover per plant.

Running a likelihood ratio test to check for the interregional independence assumption (formally, this is testing whether the dissimilarity parameter is equal to 1, and a rejection of the null hypothesis would indicate that nested logit is advised rather than conditional logit), I found that a nested logit model is preferred over conditional logit for the cases of all plants in entire

Germany, as well as for the group of small- and medium-sized plants in entire Germany. For the other groups in entire Germany, or in East and West German regions, either the IIA assumption was valid, or the model did not achieve convergence, such that the conditional logit model was the preferred model.[18]

4.2. **Results.**

4.2.1. *Baseline Results.* Table 5 shows the results from a conditional logit regression investigating the determinants of the location choices of new manufacturing plants in entire Germany. The coefficients from conditional logit estimations are shown, as well as odds ratios and average marginal effects. Regarding the interpretation of odds ratios in a location choice setting see, for example, Becker et al. (2005), and for an interpretation in terms of marginal effects see, for example, Devereux and Griffith (1998). The odds ratio is a relative probability that relates the probability of a region to be chosen as a location to the probability of not being chosen, that is $odds = \frac{P(y=1)}{P(y=0)} = \frac{\frac{1}{402}}{\frac{401}{402}} = 0.00249$.

The odds ratios that are shown in the following Tables have been recomputed to yield the effect for a 10 percent increase in the explanatory variable.[19] The marginal effects have been computed following Cameron and Trivedi (2010).[20] The increase in the location probability (average marginal effect) has also been recomputed to yield the effect for a 10 percent increase in the explanatory variable.[21] Table 6 shows the results for small- and medium-sized plants. Table 7 provides the results for large plants. Tables 8, 9, and 10 display results for East German regions and Tables 11, 12, and 13 offer results for West German regions. Results from a nested logit approach are displayed in Table 14.

Column 1 in Table 5 shows that significant effects for the location decision of new plants result from the degree of plant agglomeration, the size of the regional population, GDP growth, and

[18]The results are available from the author upon request.

[19]Because most of the explanatory variables are in logs, the odds ratio that is given by Stata would display an increase in the odds due to a 1-unit change in the log of the explanatory variable. The interpretation of these effects would be rather unintuitive. What we would rather like to talk about is the percent increase in the original (not logged) variable. To obtain the increase in odds due to a 10 percent change in the explanatory variable, one has to take the coefficient (from column 1), multiply it by ln(1.1), and then take the exponential of that term.

[20]The average marginal effects are computed according to the change of an explanatory in a baseline region. The impact on the baseline region itself as well as on all the alternative regions (i.e. cross-relationships) are computed given the methodology as described by Cameron and Trivedi. For my analyses, I chose the district *Nordsachsen* in Saxony to be the baseline region for the analyses for the entirety of Germany as well as for East Germany, and I chose the district *Dithmarschen* in Schleswig-Holstein in the North of Germany for the regressions for West Germany. Results for other baseline regions and their cross-relationships are available from the author upon request.

[21]This is done by taking the average marginal effect as given by Stata and multiplying it by ln(1.1).

the accessibility of the region measured as travel time on roads to the three nearest agglomeration centers. The odds ratios and average marginal effects allow for an interpretation of the size of impact. The results in column 2 show that a 10 percent increase in plant agglomeration in a region leads to an increase in the odds of a new plant to decide to locate in a region by 12 percent. From column 3, we can see that this translates into an increase in the probability to choose a location (baseline region is *Nordsachsen*) by 0.048 percentage points. This effect is also economically significant and important in terms of size, a fact that becomes clear when considering an example.[22] Let's say 1 plant locates in the region. Then, the baseline probability would be 1 divided by the total number of new plants, i.e. $1/1,721 = 0.000581 = p$. Consequently, an increase of 0.048 percentage points (i.e. 0.00048) is a substantial increase in the probability to locate in a region. It is easy to see that the higher p is, that is if more than 1 plant locates in a region, the smaller will be the percent increase for the location probability.[23] How much is an increase of 10 percent in the explanatory variable 'plant agglomeration'? If a region had a number of 0.005 plants per capita, and 0.0005 more plants per capita were installed (that is, if a region initially had for example 500 plants and 100000 inhabitants and 50 more plants were built in that region, resulting in a total of 550 plants), then this would increase the probability of the location choice for that region by 0.048 percentage points.

The effect for the population size shows that a 10 percent increase in the regional population increases the odds to locate in a region by 8.7 percent. In other words, if a NUTS-III region had a population size of 200000 people, an increase of 20000 people would increase the location probability of plants for this region by 0.0347 percentage points.

A decrease in average travel time by 10 percent would increase the odds of a plant's location decision by approximately 4 percent. In other words, if it took 100 minutes on average to reach the three nearest agglomeration centers on roads, a decrease by 10 more minutes of travel time on average would increase the probability of a new plant to locate in that region by about 0.0176 percentage points. The results for the travel time make clear that the plant location decision reacts in a highly sensitive manner to the regional road infrastructure. Different factors play in there: on the one hand, size and quality of roads matter. Is it easy to travel on the roads and

[22]The German Federal Statistical Office and the Offices of the Laender do not release the frequencies for a small number of plants because it would be possible to infer which plant is considered. For example, 1 plant in the region Wolfsburg might indicate that the firm Volkswagen (VW) is detected.

[23]Imagine that 7 plants are locating in the region, that is p will become 7 / 1,721 = 0.004067. An increase of 0.00048 / p amounts then to 11.8 percent.

can it be done quickly or only slowly? Are there traffic jams? Are the roads damaged? Are they sufficiently large and well-maintained for the amount of traffic? Moreover, the distance to agglomeration centers is important and it reveals whether the region is close to a network of other strong economic regions. To give an example: the average travel time to the three nearest agglomeration centers is 45 minutes for the city of Essen in the West of Germany, and 155 minutes for Flensburg in the far North close to the Danish border.[24]

Results from column 1 in Table 5 indicate that a 1 percent decrease in GDP growth leads to a 2 percent increase in the odds to locate in a region. Hence, a 1 percent decrease in GDP growth would increase the location probability for the region by 0.009 percentage points. This might indicate that an investor reacts in a forward-looking manner and is seeing potential and capacity for him to sell.

Table 6 shows that in the case of small- and medium-sized plants the effects regarding plant agglomeration, population, GDP growth, and travel time are significant. The effects are comparable in terms of economic size to the results for all plants from Table 5.

Results for large plants are displayed in Table 7. The effects show that a 10 percent increase in population leads to an increase in the odds to locate in a region of 6.3 percent. A reduction in travel time by 10 percent leads to an increase in the odds to locate in a region by 6.6 percent. The increases in the location probabilities are 0.0179 and 0.0199 percentage points, respectively.

4.2.2. *Results for East German regions.* A more refined picture can be gained from examining the location decision separately both for the East and West German regions. One will get a better idea of what made plants choose one East German region over another East German one, and one West German region over another West German one. As the results from Table 8 show, for East German regions, plant agglomeration, population, travel time, and structural funding bear significant impact on the location decision. For a 10 percent increase in plant agglomeration, the odds to locate in a region increase by 7.6 percent. A 10 percent increase in population increases the odds to locate in a region by 6 percent. The average marginal effect is 0.1012 percentage points.

A negative impact stems from the travel time on roads to reach the three nearest agglomeration centers. A 10 percent decrease in travel time leads to a 7.6 percent increase in the odds to locate

[24] Among the agglomeration centers, not only German but also neighbouring foreign countries' ones are considered for the given infrastructure variable. This information is based on the accessibility model (*Erreichbarkeitsmodell*) by the BBSR. The variable is directly available within the INKAR/ BBSR dataset.

in a region. The effect is stronger than for the entirety of Germany. Apparently, infrastructure in terms of road accessibility is an important factor for plants choosing a location among East German regions.

Moreover, regional structural funding per capita is positively significant. A 10 percent increase in structural funding per capita is increasing the odds to locate in the region by 1.8 percent. Hence, if a NUTS-III region initially had 10000 people and received 4000000 euros in terms of structural funding for infrastructure projects over the past ten years, an additional 400000 euros spent would increase the probability of a plant to locate in that region by 0.0305 percentage points. This translates into a 13.3 percent increase in the location probability for the region, assuming a baseline probability of p = 1/437= 0.0023. This effect is important in terms of economic size and reveals that regional structural funding plays an important role in plant location choices between East German regions.

Results for small- and medium-sized plants are shown in Table 9. The results are comparable to all plants, as given in Table 8. Positive and significant impacts for the location choice stem from plant agglomeration, population, and regional funding and a negative impact results from the regional road infrastructure.

Table 10 shows results for the location decision of large plants in East German regions. A positive impact stems from plant agglomeration. A 10 percent increase in plant agglomeration increases the odds to locate by approximately 21 percent.

Importantly, taxes are found to negatively impact the location decision for large plants among East German regions. A 10 percent decrease in business tax revenue per capita increases the odds of a plant to locate in the region by 13.7 percent. Thus, there is an increase in the probability of a plant to locate in a region of 0.24 percentage points or, given an assumed baseline probability p=1/20, by 4.9 percent.

A 10 percent decrease in travel time increases the odds to locate in a region by 26.5 percent. It increases the location probability by 0.51 percentage points or 10 percent, given the assumed baseline probability p. A 10 percent increase in regional funding increases the odds to locate in a region by 8.3 percent. It increases the location probability by 13.2 percentage points, which is an increase of 2.6 percent of the location probability, given the assumed baseline probability p.

4.2.3. *Results for West German regions.* Results from the analyses for West Germany are displayed for all new plants in Table 11. Agglomeration economies and the population size bear a positive impact on the plant's location decision. The odds to locate in a region (baseline region is *Dithmarschen*) increase by 14.7 and 10.5 percent, respectively, given a 10 percent increase in the variables. A negative impact stems from travel time on roads to reach the three nearest agglomeration centers. A 10 percent decrease in travel time leads to an increase in the odds to locate in the region by 3.8 percent. Results for small- and medium-sized plants are displayed in Table 12. The results are comparable to those from Table 11 for the sample of all plants. Results from Table 13 for large plants in West Germany show that only population exerts a significant impact on location decisions. A 10 percent increase in population leads to an increase in the odds to locate in a West German region by 7.6 percent.

4.2.4. *Robustness Analyses.* Nested logit results are shown in Table 14. The regression includes variables on plant agglomeration, population, GDP, GDP growth, labor costs, taxes, accessibility, regional funding and controls for plant, industry and regional fixed effects. The results reveal significant impacts for the location decision emerging from plant agglomeration, population size, travel time to the three nearest agglomeration centers, and structural funding. The coefficients from columns (1) and (3) are comparable to results from previous conditional logit estimations, in particular for plant agglomeration, population, travel time, and regional funding. As can be seen, these findings support the results from the conditional logit models. The results for small- and medium-sized plants are comparable to the sample of all plants. The dissimilarity parameters show that for the first and third groups, the random utility model condition is not fulfilled, a result that is not unusual in nested logit estimations (this is explained by Cameron and Trivedi, 2010). However, the grouping of regions, implying an investor's decision between the 25 largest German cities, South Germany, East Germany, and North-West Germany appears plausible, the parameters do not deviate much from the value 1, and the coefficients are comparable to those from the conditional logit models. In summary, plant agglomeration, the regional population, regional road infrastructure, and regional structural funding are found to economically impact the location choice of new manufacturing plants in the German economy.

5. Conclusion

In this paper, I investigate the location decisions of new manufacturing plants in the German economy and disentangle their regional determinants. New manufacturing plants base their decision to locate across the German district-free cities and districts based on a variety of factors. First, regional road infrastructure and the accessibility of a region measured by the average travel time to the three nearest agglomeration centers on roads has an important impact on the location decision. A decrease in average travel time by 10 percent is found to increase the odds of plants' location decisions for the region by 4 percent in the entirety of Germany and up to 26 percent for large plants in East German regions. This constitutes an important policy implication, given that a higher return for road infrastructure investment apparently is shown to be existent for East German regions.

Second, the amount of regional structural funding measured by the long-term spending for regional infrastructure was found to have a strong, positive influence on the decision to locate between different East German regions. East German regions offering a higher regional structural funding for infrastructure tasks per capita are apparently preferred over other East German regions. A 10 percent increase in structural funding per capita increases the odds to locate in the region by 2 percent for all plants and up to 8 percent for large plants.

Third, agglomeration economies as measured by the agglomeration of plants per region play an important role in a plant's location choice. The results reveal that a 10 percent increase in plant agglomeration increases the odds of plant location in the region by 12 percent in entire Germany and up to 21 percent for large plants in East German regions.

Fourth, a higher population in a region increases the odds of a plant to locate there. A 10 percent increase in population increases the odds to locate in that region by 9 percent in the entirety of Germany and up to 11 percent for small- and medium-sized plants in West German regions.

Policy implications arise if one wants to support the location choice of a plant in favor of one region. On the one hand, the results point to the need to reap positive benefits from supporting the agglomeration of firms and people and arising agglomeration externalities. Moreover, and importantly, the implementation and maintenance of infrastructure (particularly of the road network) appears to be an important factor for the location decision of manufacturing plants in the German economy. Most of Germany's passenger and goods transport is undertaken by road traffic. Policy measures should, however, also deal with current, pressing environmental issues. Generational sustainability and regional accessibility ought to go hand in hand. Further consideration might be given to improvements in railway transport and its regional accessibility or electric vehicles, and it would be advisable to combine the request for accessibility with climate targets. Furthermore, regional structural funding has been found to be an important factor for the location decision between East German regions, demonstrating the importance of that policy instrument for regional economic development in Germany. Support to redeem regional disparities should be sensibly targeted.

With the location choice model and the German official firm data, many more research questions could be addressed in the future, depending on the availability of further external data. It would be interesting to analyze the impact of digital infrastructure conditions (e.g. broadband internet) for the location choices of plants across the German regions.

References

Arauzo-Carod, J. M., Manjon Antolin, M. C. (2004). Firm size and geographical aggregation: An empirical appraisal in industrial location, *Small Business Economics* 22, 299-312.

Arauzo-Carod, J. M., Liviano-Solis, D., Manjon-Antolin, M. (2010). Empirical studies in industrial location: An assessment of their methods and results, *Journal of Regional Science*, 50, 3, 685–711.

Bartik, T. J. (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, 1, 14-22.

BBSR, Bundesinstitut fuer Bau-, Stadt- und Raumforschung im Bundesamt fuer Bauwesen und Raumordnung (2019). Methodische Weiterentwicklungen der Erreichbarkeitsanalysen des BBSR, BBSR Online Publikation 09/2019, Bonn.

Becker, S. O., Ekholm, K., Jaeckle, R., Muendler, M. A. (2005). Location Choice and Employment Decisions: A Comparison of German and Swedish Multinationals, *Review of World Economics* 141, 4, 693-731.

BMWI, German Federal Ministry for Economics and Industry (2019). Nationale Industriestrategie 2030 - Strategische Leitlinien fuer eine deutsche und europaeische Industriepolitik, BMWI, Berlin, www.bmwi.de.

Buch, C. M., Kleinert, J., Lipponer, A., Toubal, F. (2005). Determinants and effects of foreign direct investment: evidence from German firm-level data, *Economic Policy* 51-110.

Burda, M., Severgnini, B. (2018). Total factor productivity convergence in German states since reunification: Evidence and explanations, *Journal of Comparative Economics*, 46, 192-211.

Burda, M. C., Hunt, J. (2001). From reunification to economic integration: productivity and the labor market in Eastern Germany, *Brookings Papers on Economic Activity*, 32, 2, 1-92.

Cameron, A. C., Trivedi, P. K. (2010). Microeconometrics using Stata, Stata Press.

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, 3, 440-449.

Devereux, M. P., Griffith, R. (1998). Taxes and the location of production: evidence from a panel of US multinationals, *Journal of Public Economics* 68, 3, 335-367.

DIW Berlin (2018). Verkehr in Zahlen, Hrsg.: Bundesministerium fuer Verkehr und digitale Infrastruktur, Berlin.

Expertenkommission "Staerkung von Investitionen in Deutschland" (2015). Abschlussbericht, BMWI, PRpetuum GmBH Muenchen.

Gayko, A. (2000). Investitions- und Standortpolitik der DDR an der Oder-Neisse-Grenze 1950-1970, Peter Lang Verlag, Frankfurt.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., Shleifer, A. (1992). Growth in Cities, *Journal of Political Economy* 100, 6, 1126-1152.

Glaeser, E. L., Kerr, W. R. (2009). Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *Journal of Economics and Management Strategy*, 18, 3, 623–663.

Guimaraes, P., Figueiredo, O., Woodward, D. (2000). Agglomeration and the Location of Foreign Direct Investment in Portugal, *Journal of Urban Economics* 47, 115-135.

Hansen, E. R. (1987). Industrial Location Choice in Sao Paulo, Brazil, *Regional Science and Urban Economics* 17, 89-108.

Head, K., Ries, J., Swenson, D. (1995). Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investment in the United States, *Journal of International Economics*, 38, 223-247.

Heiss, F. (2002). Structural Choice Analysis with nested logit models, *The Stata Journal*, 2, 227-252.

Henderson, J. V. (1994). Where does an Industry Locate?, *Journal of Urban Economics*, 35, 83-104.

IWH, Leibniz-Institut fuer Wirtschaftsforschung Halle (2018). Expertisen zur Evaluation der Gemeinschaftsaufgabe "Verbesserung der regionalen Wirtschaftsstruktur" (GRW) in Sachsen-Anhalt, IWH Online, 2/2018, Halle/Saale.

Jacobs, J. (1969). The Economy of Cities, New York, Vintage Books.

Krugman, P. (1991). Geography and Trade, MIT Press, Cambridge, MA.

Marshall, A. (1890). Principles of Economics, London, MacMillan.

McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior, in Paul Zarembka (ed.), Frontiers in Econometrics, New York: Academic Press, 105-142.

Porter, M. E. (1990). The competitive advantage of nations, Harvard Business Review.

Procher, V. (2011), Agglomeration effects and the location of FDI: evidence from French first-time movers, *Annals of Regional Science*, 46, 295-312.

Roberts, M J, Tybout, J R (1997). The decision to export in Colombia: An empirical model of entry with sunk costs, *American Economic Review* 87, 4, 545-564.

Rosenthal, S. S., Strange, W. C. (2003). Geography, Industrial Organization, and Agglomeration, *Review of Economics and Statistics* 85, 2, 377-393.

Sleifer, J. (2006). Planning Ahead and Falling Behind - The East German Economy in Comparison with West Germany 1936-2002, Jahrbuch fuer Wirtschaftsgeschichte, Beiheft 8, Akademie Verlag, Berlin.

Smith, D. F., Florida, R. (1994). Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automative-Related Industries, *Journal of Urban Economics* 36, 23-41.

Appendix

Table 1: Description of variables

Variable	Description and Measurement	Data
Plant Agglomeration	Number of plants in manufacturing in a NUTS-III region, per population, logged measure	Regional Database GENESIS, based on *Unternehmensregister*, URS
Population	Number of population in a NUTS-III region, logged measure	Regional Database GENESIS
GDP	GDP per population in thousand euros, logged measure	INKAR / BBSR based on *Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Laender*, Eurostat Regio Database
GDP growth	GDP growth in relation to past year, in percent	INKAR / BBSR based on *Arbeitskreis Volkswirtschaftliche Gesamtrechnung der Laender*, Eurostat Regio Database
Labor costs	Gross wages in manufacturing per month per employee in euros, logged measure	INKAR / BBSR based on *Monats- und Jahresbericht fuer Betriebe im Bereich Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden*
Taxes	Business tax revenue per population in euros, logged measure	INKAR / BBSR based on *Realsteuervergleich des Bundes und der Laender*
Travel time/ Accessibility	Average travel time by car to the three nearest agglomeration centers, in minutes, logged measure	INKAR / BBSR based on *Erreichbarkeitsmodell by BBSR*
Structural funding	Structural regional funding of *Gemeinschaftsaufgabe zur Verbesserung der Regionalen Wirtschaftsstruktur* (GRW) for long-term infrastructure (over past ten years) per population, logged measure	INKAR / BBSR based on Database *Raumwirksame Mittel by BBSR*

Table 2: Descriptive Statistics for entire Germany

Variable	Mean	Std. Dev.	p(1)	p(99)	Obs.
Per NUTS III region ($*$ variables logged) :					
Plant Agglomeration*	-5.69	0.33	-6.51	-4.97	597242
Population*	12.03	0.64	10.59	13.92	597242
GDP*	10.27	0.31	9.72	11.20	597242
GDP growth	1.95	2.89	-6.8	9.6	597242
Labor costs*	8.08	0.21	7.62	8.60	597242
Taxes*	6.0	0.49	4.89	7.26	597242
Travel time/ Accessibility*	4.54	0.28	3.87	5.07	597242
Structural funding*	2.25	2.69	0	7.22	597313
Per NUTS III region (original data):					
Number of manufacturing plants	702.62	540.28	124	2458	597242
Population	215082.3	241194	39684	1112675	597242
GDP	30.41	11.17	16.7	73.3	597242
Labor costs	3309.17	701.9	2037.6	5446.3	597242
Taxes	456.38	265.47	132.8	1420	597242
Travel time/ Accessibility	97.57	26.08	48	159	597242
Structural funding	154.99	302.40	1	1371.4	597313

Note: The Table displays descriptive statistics for regional factors for entire Germany. Due to confidentiality restrictions regarding firm results no minimum and maximum can be shown, instead the 1st and 99th percentiles are displayed.

Table 3: Descriptive Statistics for East German regions

Variable	Mean	Std. Dev.	p(1)	p(99)	Obs.
Per NUTS III region (∗ variables logged):					
Plant Agglomeration*	-5.79	0.34	-6.52	-4.97	129075
Population*	11.93	0.65	10.49	15.03	129075
GDP*	10.04	0.18	9.72	10.46	129075
GDP growth	1.95	3.25	-6.8	11.8	129075
Labor costs*	7.84	0.15	7.59	8.26	129075
Taxes*	5.58	0.29	4.89	6.27	129075
Travel time/ Accessibility*	4.72	0.19	4.13	5.11	129075
Structural funding*	6.29	0.60	3.87	7.49	129075
Per NUTS III region (original data):					
Number of manufacturing plants	597.56	665.56	87	5504	129075
Population	209892.9	379572.4	35967	3375222	129075
GDP	23.27	4.43	16.7	34.9	129075
Labor costs	2559.0	394.04	1969.4	3848.9	129075
Taxes	277.94	85.55	132.8	527.6	129075
Travel time/ Accessibility	113.56	20.49	62	166	129075
Structural funding	632.45	349.55	48	1793.9	129075

Note: The Table displays descriptive statistics for regional factors for East German regions. Due to confidentiality restrictions regarding firm results no minimum and maximum can be shown, instead the 1st and 99th percentiles are displayed.

Table 4: Descriptive Statistics for West German regions

Variable	Mean	Std. Dev.	p(1)	p(99)	Obs.
Per NUTS III region (∗ variables logged):					
Plant Agglomeration*	-5.67	0.33	-6.42	-4.96	468112
Population*	12.06	0.64	10.59	13.92	468112
GDP*	10.33	0.31	9.7	11.29	468112
GDP growth	1.95	2.78	-7.2	8.7	468112
Labor costs*	8.15	0.17	7.81	8.62	468112
Taxes*	6.11	0.47	4.79	7.51	468112
Travel time/ Accessibility*	4.5	0.28	3.85	5.07	468112
Structural funding*	1.13	1.84	0	5.55	468112
Per NUTS III region (original data):					
Number of manufacturing plants	731.64	496.36	131	2458	468112
Population	216525.9	185700.1	39684	1112675	468112
GDP	32.37	11.65	16.3	80	468112
Labor costs	3516	622.66	2463.7	5529.7	468112
Taxes	505.57	276.87	120.6	1828.6	468112
Travel time/ Accessibility	93.16	25.73	47	159	468112
Structural funding	23.38	52.99	1	256	468112

Note: The Table displays descriptive statistics for regional factors for West German regions. Due to confidentiality restrictions regarding firm results no minimum and maximum can be shown, instead the 1st and 99th percentiles are displayed.

Table 5: Location choices - entire Germany, all plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	1.2127***	1.1225	0.048
	(0.0974)		
Population	0.8757***	1.087	0.0347
	(0.0426)		
GDP	0.0078	1.0007	0.0003
	(0.1798)		
GDP growth	-0.0216**	0.9786	-0.009
	(0.0089)		
Labor costs	-0.2699	0.9746	-0.0107
	(0.2504)		
Taxes	0.0470	1.0045	0.0019
	(0.1226)		
Travel time/ Accessibility	-0.4452***	0.9585	-0.0176
	(0.111)		
Structural funding	0.0228	1.0231	0.0009
	(0.0198)		
Number of observations	597,187		
Number of plants	1,721		
Wald χ^2	801.42		
Prob > χ^2	0.0000		
Log pseudolikelihood	-9715.0419		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/1,721. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 6: Location choices - entire Germany, small- and medium-sized plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	1.2743***	1.1291	0.0512
	(0.1001)		
Population	0.8931***	1.0888	0.0358
	(0.0442)		
GDP	-0.0578	0.9945	-0.0023
	(0.1908)		
GDP growth	-0.0251***	0.9752	-0.0106
	(0.0091)		
Labor costs	-0.2132	1.0205	-0.0085
	(0.2608)		
Taxes	0.062	1.0059	0.0025
	(0.1310)		
Travel time/ Accessibility	-0.4171***	0.9610	-0.0167
	(0.1154)		
Structural funding	0.0182	1.0017	0.0007
	(0.0208)		
Number of observations	558,323		
Number of plants	1,609		
Wald χ^2	797.03		
Prob > χ^2	0.0000		
Log pseudolikelihood	-9064.2781		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/1,609. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 7: Location choices - entire Germany, large plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	0.2577	1.0249	0.0072
	(0.4284)		
Population	0.6438***	1.0633	0.0179
	(0.1619)		
GDP	0.6909	1.0681	0.0192
	(0.4968)		
GDP growth	0.0313	1.0318	0.0091
	(0.0413)		
Labor costs	-1.0316	0.9064	-0.0287
	(0.8821)		
Taxes	-0.0797	0.9924	-0.0022
	(0.2924)		
Travel time/ Accessibility	-0.7173*	0.9339	-0.0199
	(0.4144)		
Structural funding	0.0494	1.0047	0.0014
	(0.0684)		
Number of observations	38,864		
Number of plants	112		
Wald χ^2	32.56		
Prob > χ^2	0.0000		
Log pseudolikelihood	-640.03535		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/112. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 8: Location choices - East Germany, all plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	0.7638***	1.0755	0.1268
	(0.1798)		
Population	0.6097***	1.0598	0.1012
	(0.0796)		
GDP	0.0887	1.0085	0.0146
	(0.4805)		
GDP growth	-0.024	0.9763	-0.0417
	(0.0167)		
Labor costs	-0.6445	0.9404	-0.1069
	(0.4642)		
Taxes	-0.1272	0.9879	-0.0211
	(0.2699)		
Travel time/ Accessibility	-0.8399***	0.9231	-0.1392
	(0.3131)		
Structural funding	0.1839**	1.0177	0.0305
	(0.0863)		
Number of observations	32,775		
Number of plants	437		
Wald χ^2	117.06		
Prob > χ^2	0.0000		
Log pseudolikelihood	-1830.8992		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/437. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 9: Location choices - East Germany, small- and medium-sized plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	0.7135***	1.0704	0.1182
	(0.1846)		
Population	0.6268***	1.0616	0.1038
	(0.0808)		
GDP	-0.0091	0.9991	-0.0015
	(0.5005)		
GDP growth	-0.0287*	0.9717	-0.0499
	(0.017)		
Labor costs	-0.6763	0.9376	-0.1118
	(0.4778)		
Taxes	-0.0604	0.9943	-0.01
	(0.2818)		
Travel time/ Accessibility	-0.7570**	0.9304	-0.1252
	(0.3191)		
Structural funding	0.1575*	1.0151	0.0261
	(0.0869)		
Number of observations	31,275		
Number of plants	417		
Wald χ^2	114.28		
Prob > χ^2	0.0000		
Log pseudolikelihood	-1746.2168		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/417. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 10: Location choices - East Germany, large plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	1.9952**	1.2094	0.3155
	(0.8054)		
Population	0.1063	1.0102	0.0168
	(0.6478)		
GDP	2.1636	1.229	0.3417
	(1.55)		
GDP growth	0.0528	1.0542	0.0875
	(0.077)		
Labor costs	-0.4893	0.9544	-0.0773
	(1.9877)		
Taxes	-1.5470**	0.8629	-0.2444
	(0.7544)		
Travel time/ Accessibility	-3.2362*	0.7346	-0.5109
	(1.896)		
Structural funding	0.8366*	1.083	0.1322
	(0.4792)		
Number of observations	1,500		
Number of plants	20		
Wald χ^2	18.91		
Prob > χ^2	0.0153		
Log pseudolikelihood	-80.758123		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/20. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 11: Location choices - West Germany, all plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	1.4396***	1.1471	0.0130
	(0.1191)		
Population	1.0517***	1.1054	0.0095
	(0.0585)		
GDP	-0.0808	0.9923	-0.0007
	(0.2041)		
GDP growth	-0.0221**	0.9782	-0.0021
	(0.0104)		
Labor costs	-0.1090	0.9897	-0.001
	(0.2919)		
Taxes	0.0358	1.0034	0.0003
	(0.1420)		
Travel time/ Accessibility	-0.4105***	0.9616	-0.0037
	(0.1208)		
Structural funding	0.0355	1.0034	0.0003
	(0.0218)		
Number of observations	349,248		
Number of plants	1,284		
Wald χ^2	576.02		
Prob > χ^2	0.0000		
Log pseudolikelihood	-6887.7819		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Dithmarschen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/1,284. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 12: Location choices - West Germany, small- and medium-sized plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	1.5620***	1.1605	0.0131
	(0.1218)		
Population	1.0776***	1.1082	0.0091
	(0.0617)		
GDP	-0.1459	0.9862	-0.0012
	(0.2181)		
GDP growth	-0.0253**	0.975	-0.0022
	(0.0107)		
Labor costs	-0.0202	0.9981	-0.0002
	(0.3052)		
Taxes	0.0433	1.0041	0.0004
	(0.1532)		
Travel time/ Accessibility	-0.4036***	0.9623	-0.0034
	(0.1258)		
Structural funding	0.0372	1.0036	0.0003
	(0.023)		
Number of observations	324,224		
Number of plants	1,192		
Wald χ^2	565.19		
Prob > χ^2	0.0000		
Log pseudolikelihood	-6375.2575		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Dithmarschen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/1,192. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 13: Location choices - West Germany, large plants

	(1)	(2)	(3)
	Conditional Logit	Odds Ratio	Average Marginal Effects in % points
Plant Agglomeration	-0.2082	0.9804	-0.0046
	(0.514)		
Population	0.7697***	1.0761	0.0168
	(0.1840)		
GDP	0.5313	1.0519	0.0116
	(0.5273)		
GDP growth	0.0238	1.0240	0.0055
	(0.0432)		
Labor costs	-1.245	0.8881	-0.0273
	(0.9850)		
Taxes	-0.0264	0.9975	-0.0006
	(0.3072)		
Travel time/ Accessibility	-0.3837	0.9641	-0.0084
	(0.4468)		
Structural funding	-0.0048	0.9995	-0.0001
	(0.0731)		
Number of observations	25,024		
Number of plants	92		
Wald χ^2	38.96		
Prob > χ^2	0.0000		
Log pseudolikelihood	-500.72647		

Note: This Table displays estimates for the location choice of new German manufacturing plants using a conditional logit model. The coefficients in columns (2) and (3) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Dithmarschen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/92. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Robust standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 14: Robustness checks - location choices - nested logit model

	(1)	(2)	(3)	(4)
	Entire Germany			
	All plants		Small and	medium-sized plants
	Nested Logit	Average Marginal Effects in % points	Nested Logit	Average Marginal Effects in % points
Explanatory factors				
Plant Agglomeration	0.8615**	0.0248	1.0897**	0.0261
	(0.4375)		(0.4829)	
Population	0.8340**	0.0240	1.063**	0.0254
	(0.3983)		(0.446)	
GDP	0.0220	0.0006	-0.0050	-0.0001
	(0.1462)		(0.1844)	
GDP growth	-0.0148	-0.0045	-0.0196	-0.0049
	(0.0104)		(0.0125)	
Labor costs	-0.2085	-0.0060	-0.292	-0.007
	(0.2255)		(0.2823)	
Taxes	0.0063	0.0002	0.0113	0.0003
	(0.0899)		(0.1135)	
Travel time/ Accessibility	-0.3644*	-0.0105	-0.4299*	-0.0103
	(0.213)		(0.2305)	
Structural funding	0.0464*	0.0013	0.0571*	0.0014
	(0.027)		(0.0307)	
Dissimilarity parameters (τ_k)				
25 largest German cities	1.2043		1.5713	
	(0.4829)		(0.5658)	
South German district-free cities and districts	0.7023		0.8739	
	(0.3504)		(0.3791)	
East German district-free cities and districts	1.3461		1.6479	
	(0.5858)		(0.609)	
North-West German district-free cities and districts	0.6425		0.7948	
	(0.3204)		(0.3493)	
LR test for IIA (τ_k=1)	59.73		65.23	
Prob > χ^2	0.0000		0.0000	
Number of observations	540,973		530,910	
Number of plants	1,559		1,530	
Wald χ^2	36.96		32.57	
Prob > χ^2	0.0007		0.0033	
Log likelihood	-8748.5987		-8572.9989	

Note: This Table displays estimates for the location choice of new German manufacturing plants using a nested logit model. The coefficients in columns (2) and (4) show the effect for a 10 percent increase in the explanatory variable, except for GDP growth which covers a 1 percent increase. Average marginal effects are computed according to Cameron and Trivedi (2010). The baseline region is chosen to be the district *Nordsachsen*. Assuming that 1 plant locates in this region, the baseline location probability would be p=1/ number of plants = 1/1,559 for the sample of all plants and p=1/1,530 for the sample of small- and medium-sized plants. The dependent variable is a dummy variable for each plant and region. The variable is equal to 1 for a chosen district-free city or district and 0 for the other regions. Plant-level, industry-level and regional specific fixed effects are controlled for. Data are taken from the German Federal Statistical Office and the Offices of the Laender, from the Regional Database GENESIS and from INKAR/ BBSR. Standard errors were computed and are displayed in parentheses. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 15: Description of nests

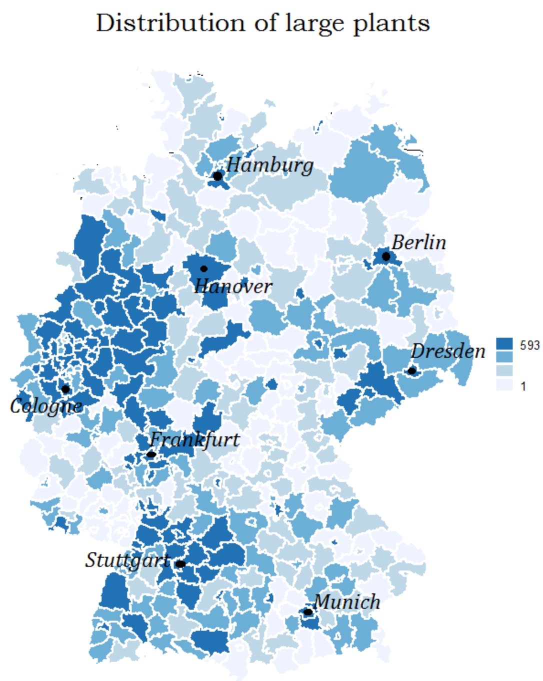
25 largest German cities	Berlin, Hamburg, Munich, Cologne, Frankfurt am Main, Stuttgart, Duesseldorf, Leipzig, Dortmund, Essen, Bremen, Dresden, Hannover, Nuernburg, Duisburg, Bochum, Wuppertal, Bielefeld, Bonn, Muenster, Karlsruhe, Mannheim, Augsburg, Wiesbaden, Moenchengladbach
South German district-free cities and districts (in Bavaria, Baden-Wuertemberg)	Amberg, Ansbach, Aschaffenburg, Baden-Baden, Bamberg, Bayreuth, Coburg, Erlangen, Freiburg, Fuerth, Heidelberg, Heilbronn, Hof, Ingolstadt, Kaufbeuren, Kempten, Landshut, Memmingen, Passau, Pforzheim, Regensburg, Rosenheim, Schwabach, Schweinfurt, Straubing, Ulm, Weiden, Wuerzburg, Aichach-Friedberg, Alb-Donau-Kreis, Altoetting, Amberg-Sulzbach, Ansbach, Aschaffenburg, Augsburg, Bad Kissingen, Bad Toelz-Wolfratshausen, Bamberg, Bayreuth, Berchtesgadener Land, Biberach, Boeblingen, Bodenseekreis, Breisgau-Hochschwarzwald, Calw, Cham, Coburg, Dachau, Deggendorf, Dillingen an der Donau, Dingolfing-Landau, Donau-Ries, Ebersberg, Eichstaett, Emmendingen, Enzkreis, Erding, Erlangen-Hoechstadt, Esslingen, Forchheim, Freising, Freudenstadt, Freyung-Grafenau, Fuerstenfeldbruck, Fuerth, Garmisch-Partenkirchen, Goeppingen, Guenzburg, Hassberge, Heidenheim, Heilbronn, Hof, Hohenlohekreis, Karlsruhe, Kelheim, Kitzingen, Konstanz, Kronach, Kulmbach, Landsberg am Lech, Landshut, Lichtenfels, Lindau, Loerrach, Ludwigsburg, Main-Spessart, Main-Tauber-Kreis, Miesbach, Miltenberg, Muehldorf am Inn, Muenchen, Neckar-Odenwald-Kreis, Neu-Ulm, Neuburg-Schrobenhausen, Neumark in der Oberpfalz, Neustadt an der Aisch-Bad Windsheim, Neustadt an der Waldnaab, Nuernberger Land, Oberallgaeu, Ortenaukreis, Ostalbkreis, Ostallgaeu, Passau, Pfaffenhofen an der Ilm, Rastatt, Ravensburg, Regen, Regensburg, Rems-Murr-Kreis, Reutlingen, Rhein-Neckar-Kreis, Rhoen-Grabfeld, Rosenheim, Roth, Rottal-Inn, Rottweil, Schwandorf, Schwarzwald-Baar-Kreis, Schwaebisch Hall, Schweinfurt, Sigmaringen, Starnberg, Straubing-Bogen, Tirschenreuth, Traunstein, Tuebingen, Tuttlingen, Unterallgaeu, Waldshut, Weilheim-Schongau, Weissenburg-Gunzenhausen, Wunsiedel im Fichtelgebirge, Wuerzburg, Zollernalbkreis
East German district-free cities and districts (in Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thuringen)	Brandenburg an der Havel, Chemnitz, Cottbus, Dessau-Rosslau, Eisenach, Erfurt, Frankfurt an der Oder, Gera, Halle an der Saale, Jena, Magdeburg, Potsdam, Rostock, Schwerin, Suhl, Weimar, Altenburger Land, Altmarkkreis Salzwedel, Anhalt-Bitterfeld, Barnim, Bautzen, Boerde, Burgenlandkreis, Dahme-Spreewald, Eichsfeld, Elbe-Elster, Erzgebirgskreis, Goerlitz, Gotha,Greiz, Harz, Havelland, Hildburghausen, Ilm-Kreis, Jerichower Land, Kyffhaeuserkreis, Leipzig, Ludwigslust-Parchim, Mansfeld-Suedharz, Maerkisch-Oderland, Mecklenburgische Seenplatte, Meissen, Mittelsachsen, Nordhausen, Nordsachsen, Nordwestmecklenburg, Oberhavel, Oberspreewald-Lausitz, Oder-Spree, Ostprignitz-Ruppin, Potsdam-Mittelmark, Prignitz, Rostock, Saale-Holzland-Kreis, Saalekreis, Saale-Orla-Kreis, Saalfeld-Rudolstadt, Saechsische Schweiz-Osterzgebirge, Salzlandkreis, Schmalkalden-Meiningen, Soemmerda, Sonneberg, Spree-Neisse, Stendal, Teltow-Flaeming, Uckermark, Unstrut-Hainich-Kreis, Vogtlandkreis, Vorpommern-Greifswald, Vorpommern-Ruegen, Wartburgkreis, Weimarer Land, Wittenberg, Zwickau

Table 15: Continued - Description of nests

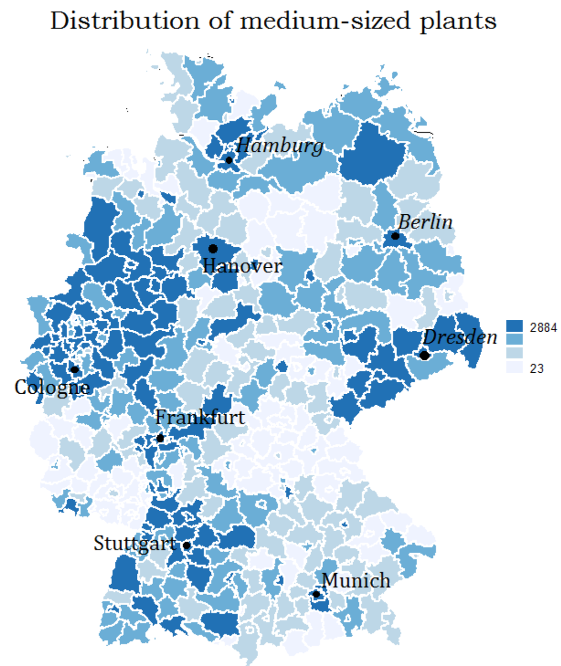
North-West German district-free cities and districts (in Schleswig-Holstein, Niedersachsen, Nordrhein-Westfalen, Hessen, Saarland, Rheinland-Pfalz)	Aachen, Bottrop, Braunschweig, Bremerhaven, Darmstadt, Delmenhorst, Emden, Flensburg, Frankenthal, Gelsenkirchen, Goettingen, Hagen, Hamm, Herne, Kaiserslautern, Kassel, Kiel, Koblenz, Krefeld, Landau, Leverkusen, Luebeck, Ludwigshafen, Mainz, Muelheim, Neumuenster, Neustadt an der Weinstrasse, Oberhausen, Offenbach, Oldenburg, Osnabrueck, Pirmasens, Remscheid, Salzgitter, Solingen, Speyer, Trier, Wilhelmshaven, Wolfsburg, Worms, Zweibruecken, Aachen, Ahrweiler, Altenkirchen, Alzey-Worms, Ammerland, Aurich, Bad Duerkheim, Bad Kreuznach, Bergstrasse, Bernkastel-Wittlich, Birkenfeld, Borken, Celle, Cloppenburg, Cochem-Zell, Coesfeld, Cuxhaven, Darmstadt-Dieburg, Diepholz, Dithmarschen, Donnersbergkreis, Dueren, Eifelkreis Bitburg-Pruem, Emsland, Ennepe-Ruhr-Kreis, Euskirchen, Friesland, Fulda, Germersheim, Giessen, Gifhorn, Goslar, Goettingen, Grafschaft Bentheim, Gross-Gerau, Guetersloh, Hameln-Pyrmont, Hannover, Harburg, Heidekreis, Heinsberg, Helmstedt, Herford, Hersfeld-Rotenburg, Herzogtum Lauenburg, Hildesheim, Hochsauerlandkreis, Hochtaunuskreis, Holzminden, Hoexter, Kaiserslautern, Kassel, Kleve, Kusel, Lahn-Dill-Kreis, Leer, Limburg-Weilburg, Lippe, Luechow-Dannenberg, Lueneburg, Main-Kinzig-Kreis, Main-Taunus-Kreis, Mainz-Bingen, Marburg-Biedenkopf, Maerkischer Kreis, Mayen-Koblenz, Merzig-Wadern, Mettmann, Minden-Luebbecke, Neunkirchen, Neuwied, Nienburg, Nordfriesland, Northeim, Oberbergischer Kreis, Odenwaldkreis, Offenbach, Oldenburg, Olpe, Osnabrueck, Osterholz, Osterode, Ostholstein, Paderborn, Peine, Pinneberg, Ploen, Recklinghausen, Rendsburg-Eckernfoerde, Rhein-Erft-Kreis, Rheingau-Taunus-Kreis, Rhein-Hunsrueck-Kreis, Rheinisch-Bergischer Kreis, Rhein-Kreis Neuss, Rhein-Lahn-Kreis, Rhein-Pfalz-Kreis, Rhein-Sieg-Kreis, Rotenburg, Saarbruecken, Saarlouis, Saarpfalz-Kreis, Schaumburg, Schleswig-Flensburg, Schwalm-Eder-Kreis, Segeberg, Siegen-Wittgenstein, Soest, St. Wendel, Stade, Steinburg, Steinfurt, Stormarn, Suedliche Weinsteinstrasse, Suedwestpfalz, Trier-Saarburg, Uelzen, Unna, Vechta, Verden, Viersen, Vogelsbergkreis, Vulkaneifel, Waldeck-Frankenberg, Warendorf, Werra-Meissner-Kreis, Wesel, Wesermarsch, Westerwaldkreis, Wetteraukreis, Wittmund, Wolfenbuettel

Note: This Table shows the composition regarding NUTS-III regions for the 4 nests used in the nested logit model. District-free cities and districts might have the same name, so they will show up twice in the nests, e.g. the district-free city Bamberg and the district Bamberg.

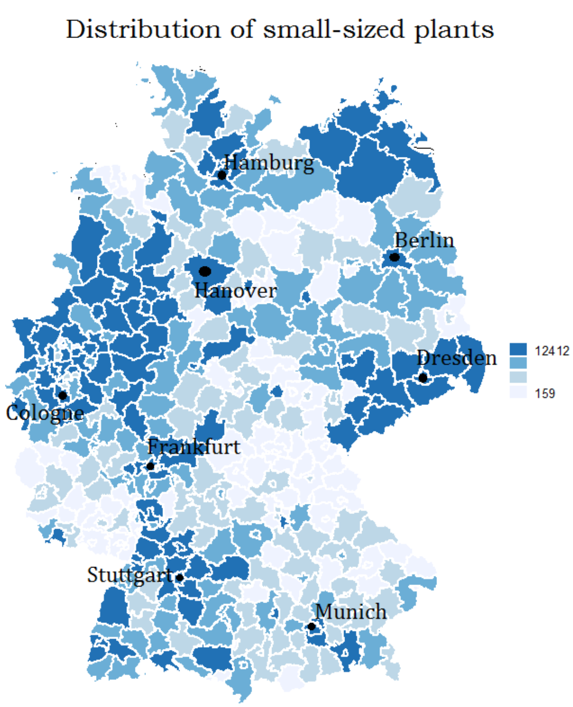
Figure 1



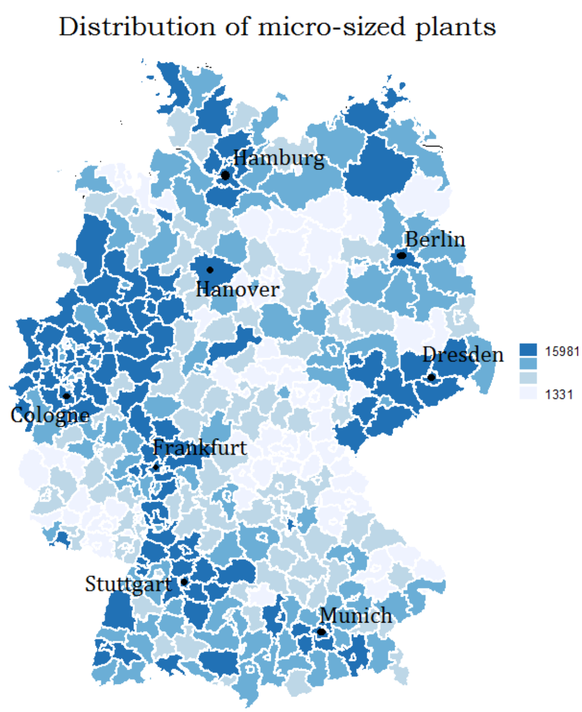
(A)



(B)



(C)



(D)

Note: This Figure shows the distribution of plants in sectors B-N and P-S across NUTS-III regions (district-free cities and districts) in 2013 for Germany. Data are taken form the Regional Database (GENESIS), based on the register of firms (URS). Micro-sized plants are defined as having 0-9 employees that are covered by social insurance, small-sized plants have 10-49 employees, medium-sized plants have 50-249 employees and large plants have 250 and more employees. The colored groups in each graphic have to be interpreted as follows: 25 percent of the regions are covered by one of the 4 colors. That is the darkest color displays the 25 percent regions with the largest number of plants and the lightest color displays the 25 percent of regions with the lowest number of plants per employment size category.