

Donges, Alexander; Streb, Jochen

Conference Paper

How the West was Settled. The Location Choice of East German Companies Migrating to West Germany after World War II

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2024: Upcoming Labor Market Challenges

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Donges, Alexander; Streb, Jochen (2024) : How the West was Settled. The Location Choice of East German Companies Migrating to West Germany after World War II, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2024: Upcoming Labor Market Challenges, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/302389>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

How the West was Settled. The Location Choice of East German Companies Migrating to West Germany after World War II*

August 11, 2023

Alexander Donges (University of Mannheim)[†]

Jochen Streb (University of Mannheim)[‡]

Abstract

The division of Germany into a free-market West and a socialist East that followed the military defeat in spring 1945 induced thousands of East German companies to relocate to the West in the following years. This exogenous mass exodus allows us to identify the motives for the choice of firm location in a natural experiment. We test whether the East German firms were primarily attracted by existing West German agglomeration economies or rather sought new locations geographically close to their original homes, which allowed them to retain their access to pre-existing local networks. To test the determinants of the firms' location choices, we use a newly constructed data set including information for over 4,200 relocated Eastern German firms, which we combine with county-level data on local economic activity and other socio-economic characteristics. By applying a mixed logit choice model, we find a negative effect of distance. Firms preferred places close to their original locations with market conditions they already knew. The fact that this negative distance effect is stronger for firms from original places close to the inner-German border strengthens our hypothesis that "home advantages" mattered. We also provide evidence for the attractiveness of agglomerations showing that firms favored places with high productivity and market potential. There are heterogeneous effects across industries: For companies in global market-oriented industries, the agglomeration effect is stronger and the location advantage smaller than for companies in industries with a more local customer or supplier base.

JEL Codes D22, N64, N94, R12

Key words agglomeration economies, German division, home bias, location choice, mixed logit choice model

* This research was supported by the *Deutsche Forschungsgemeinschaft* (DFG). We thank Ulrich Pfister, Mark Spoerer, and Nikolaus Wolf for their helpful comments and discussions. We are also grateful for valuable comments by conference participants at the Congress for Economic and Social History in Leipzig, the Conference of the Economic History Committee of the VfS in Bochum, the Baden-Württemberg Economic History Workshop in Tübingen, and by seminar participants at London School of Economics, Universidad Carlos III de Madrid, the University of Regensburg, HU Berlin, and the University of Mannheim. Thanks to Felix Kersting, who shared data from the 1925 employment census with us. We are especially grateful for the late Jennifer Köhler, who provided excellent research assistance in the early stages of this project. We also thank Henri Felshart, Conrad Hoßfeld, Lukas Trautner, and Sergej Zuev for their research assistance.

[†] E-Mail: donges@uni-mannheim.de

[‡] E-Mail: jochen.streb@uni-mannheim.de

1 Introduction

The division of Germany into a free-market West and a socialist East that followed the military defeat in spring of 1945 motivated thousands of East German companies to relocate to the West in the years that followed. We use this mass exodus as a natural experiment to identify the economic motives behind the location choices of firms. Our results go beyond the usual finding that firm owners prefer large agglomerations when choosing a location. We show that East German firms were also attracted to the benefits that accrue when they relocate in a nearby Western region whose market conditions they already know well. This “home advantage” of low transaction costs that result from already knowing customers, suppliers, or investors sometimes even exceeded the advantages of large agglomerations.

There are two main explanations why firms do not distribute evenly within a given geographic area, as proposed in the circular economy of Salop (1979), but rather cluster in particular cities and regions (Davis and Weinstein 2002). On the one hand, companies are attracted to the fundamental geographic advantages of a place, including raw material deposits or near-by navigable water ways, which in both cases promise to reduce transportation costs (Ellison and Glaeser 1999). On the other hand, firms settle in geographic proximity to potential customers and other firms from which they expect positive externalities. The agglomeration advantages of many firms from the same industry include the local availability of a skilled workforce, specialized suppliers, and a range of services tailored to the needs of the industry (Marshall 1890, Krugman 1991, Ellison et al. 2010). Proximity to competitors, customers, and suppliers also facilitates the exchange of ideas and know-how across firms and industries, resulting in more diverse innovation when compared to an isolated site (Jacobs 1970, Malmberg and Maskell 2002). Because of increasing returns to size, an upward spiral occurs in metropolitan areas. New firms are attracted to the advantages of agglomeration, adding to the positive externalities of the place, while thereby initiating even more entrepreneurial activity such as business start-ups (Capello 2019, p. 32). Using the analogy of Glaeser (2011), this process contributes to the “triumph of the city”.¹

Since the 1980s, economists use data to test the hypothesis that firms prefer agglomerations to other locations (Carlton 1983, Bartik 1985, Hansen 1987). These studies assume that companies compare the available locations in an unbiased way and select the place

¹ There might be limits to this upward spiral when rising rents and other costs of overcrowding erode the benefits of agglomeration (Glaeser 1998).

where the highest firm profit can be expected. However, Figueiredo et al (2002) provide evidence that this assumption does not hold for start-ups, showing that founders prefer their home region as a business location simply out of attachment to their family and friends. A prominent example is the software company Microsoft, which was founded in Redmond, Washington, close to Seattle, where Bill Gates grew up. The home region may offer many economic advantages. Founders can rely on established social networks at home that includes potential suppliers, customers, and investors, and they know more about the economically relevant characteristics of their home region than they do for any other region. This home bias, combined with the fact that densely populated agglomerations produce a higher number of founders for demographic reasons alone, means that national samples of business start-ups regularly include a high number of firms that were established in agglomerations for reasons unrelated to increasing returns to size. Studies that rely on national samples of business start-ups might therefore systematically overestimate the attractiveness of agglomeration economies. Using the U.S. biopharmaceutical industry as an example, Alcácer and Delgade (2016) show that already established companies may have an interest in locating new operations and branches close to existing company sites because they expect spillovers between neighboring old and new facilities. In this case of internal agglomeration, the positive effect of external agglomeration would also be overestimated.

Studies analyzing the location choices of firms on basis of modern data mainly focus on foreign direct investment (FDI) decisions. For example, Head and Mayer (2004) show that market potential mattered for the location choice of Japanese firms in the European Union, Chung and Alcácer (2002) investigate the link between local R&D intensity and FDI in the United States, Devereux and Griffith (1998) analyze the effect of taxation on investment of US multinationals in Europe, and Becker et al. (2005) study how local labor market conditions affected FDI decisions of German and Swedish multinationals. Guimarães et al (2000) observe that, in Portugal, subsidiaries of foreign firms were founded primarily in regions with a strong service sector and high industry-specific employment. Hecht (2017) investigates the location choices of German firms in the Czech Republic after the fall of the Iron Curtain, identifying a preference for densely populated areas where firms from the same industry are disproportionately represented.

Pascal and McCall (1980) cast doubt that firms make their location choices under complete information and provide evidence for “chain migration”.² In their view, firms are uncertain about many characteristics of the target regions available for selection. To save the costs of information gathering, they therefore simply set up operations where many have already done so before.³ Following this argument, the location choice can be interpreted as a form of herd behavior.⁴ Thus, the empirical observation that firms settle in agglomerations does not necessarily imply the existence of positive externalities in the sense of Marshall (1890).

All studies that examine the location choice based on FDI data have in common that they focus on a particular subset of firms that are typically more productive and larger than the national average (Helpman et al 2004). Managers of multinational companies should have more experience and expertise than a first-time founder or the manager of a small or medium-sized enterprise (SME), when deciding where to locate business activities. By contrast, the relocation of East German firms to the West in the wake of Germany’s division after World War II involved not only large firms but also many SMEs from a broad spectrum of industries, ranging from engineering and textiles to publishing, banking, insurance, and retailing. Therefore, the exogenous relocation of East German companies after WWII provides a rare opportunity to study the location decision of a more heterogeneous group of firms.

Earlier studies tried to infer the motives of East German firms’ location choice from the geographic distribution of relocated companies but lack a rigorous empirical investigation. By using the western addresses of about 5,200 relocated East German firms, Hefe (1998, p. 119) notices that these firms settled either near the inner German border (northern Hesse, greater Hanover, northern Bavaria), the so-called Zonal Border Area (*Zonenrandgebiet*),⁵ or in the later economic “powerhouses” of western Germany (along the Stuttgart-Munich axis, Rhine-Main area, Upper Rhine Valley, Ruhr area). Hefe’s (1998) second observations speaks to the general attraction effect of West German agglomerations, with East German firms perhaps guided more by herd instinct than by rational considerations of expected increasing returns to

² The fact that emigrants benefit from information they receive from their predecessors when choosing their own destination has been discussed at length in the context of the great emigration waves of the 19th century. See, for example, Wegge (1998).

³ In the flowery words of Pascal and McCall (1980, p. 386), “Adrift in a sea of uncertainty, what would be more logical than heading toward those islands which have demonstrably provided salvation to similar and early swimmers?”

⁴ On the general relationship between search costs and herd behavior, see Bikhchandai et al (1998).

⁵ Beginning in 1951, the Federal German government felt compelled to support distressed areas, including the Zonal Border Area, with loans for firms and communities. According to a resolution passed by the German Bundestag on July 20, 1953, counties and independent cities that were within 40 kilometers from the inner-German border, the Baltic Sea coast or the border with Czechoslovakia belonged to the Zonal Border Area. However, a law on zonal border development was not enacted until 1971. See Nuppenau (1974), especially p. 27 f.

size in the very uncertain postwar period. Concerning Hefe's first observation, Held (1959, pp. 330-334) had already surmised that many East German enterprises from the eastern vicinity of the inner-German border relocated to places in the western Zonal Border Area, where they expected economic and cultural conditions like those at home. Companies from the more eastern part of the later GDR would not have felt this home bias and therefore would not have shown a preference for the Zonal Border Area.

A first attempt to examine the location choices of East German firms using a discrete choice model was made by Falck et al (2013), building on research by Buenstorf and Guenther (2010). Based on a small sample of only 33 machine-tool firms, the authors sought to explain the location decision in terms of a preference for technological or cultural proximity.⁶ However, geographic distance emerges as the most important factor in their study. The firms apparently preferred a new location as far away from the inner-German border as possible, which could be interpreted as a desire to put as much distance as possible between themselves and the Soviet occupation forces that is to increase the distance to the communist threat. The observation that relocated firms preferred distant sites over nearby ones contradicts not only the theoretical assumption of a negative influence of distance due to higher transaction costs, but also the observations by Held (1959) and Hefe (1998) and the empirical results of this paper that relies on a much broader sample, including over 4,200 firms in a wide range of industries.

To investigate the motives of location choice, we apply a mixed logit choice model and test to what extent firms preferred new places close to their old sites in the East compared to established agglomerations in the West. We control for a battery of other potential socio-economic factors that may have affected the location decision, including institutional differences, local access to human capital, infrastructure, financial incentives, and similar industry structures. The results show that the home advantage mattered, which means that there is a negative effect of distance to the old firm location on location choice. Moreover, we find heterogeneous effects by comparing firms of different pre-war locations. The negative effect of distance is stronger for firms that were originally located close to the West-East border that was established in 1945 compared to firms that were from places more far away in the East. This result underlines the importance of local markets, when choosing a new firm location. We also observe heterogeneous effects across industries. For companies in global market-oriented

⁶ Technological proximity would be present if the portfolio of types of machine tools produced in a firm's region of origin and in the region of destination largely coincided. Cultural proximity is measured by Falck et al (2013) based on the similarity of German dialects spoken in the region of origin and destination region.

industries, the agglomeration effect is stronger and the location advantage smaller than for companies in industries with a more local customer or supplier base.

The entrepreneurs and executives who went to West Germany to re-start their business represented only a small fraction of the millions of people who moved to West Germany after 1945 from the Soviet occupation zone, but mainly and often involuntarily from the lost German territories east of the Oder-Neisse line, from the Sudetenland, Yugoslavia, Romania, and Hungary. In 1950, there were 7.88 million expellees from the latter regions and 1.56 million refugees from the Soviet occupation zone in West Germany. This was over twenty percent of the total population (Spoerer and Streb 2013, p. 212).

Economic research on expellees focuses on the consequences for the West German labor market (Braun and Mahmoud 2014, Braun et al. 2021) and the economic integration of immigrants (Bauer et al. 2013). Most studies emphasize that, similar to relocated East German businesses, displaced persons (initially) settled as close as possible to their place of origin. For example, while expellees from East Prussia tended to move to northern West Germany, Sudeten Germans displaced from Czechoslovakia settled mainly in Bavaria. Semrad (2015) suggests that the well-educated Sudeten Germans were instrumental in expanding secondary education in their new homeland and thus contributed to economic growth in Bavaria. Schuman (2014), Wyrwich (2020), and Ciccone and Nimczik (2022) examine the economic effects of the ban on immigration in the French occupation zone and provide evidence that it harmed long-term economic development in these regions. In line with this argument, Peters (2022) provides empirical evidence for a positive effect of refugee settlement on economic development, showing that it increased income per capita and triggered industrialization in rural areas.

If the assumption is true that agglomeration is caused by increasing returns to size, then as far as the spatial distribution of economic activities is concerned, there exists no unique optimum that is determined by fundamental factors. Rather, it depends on historical coincidences which of the many conceivable equilibria is realized. To prove the empirical existence of multiple equilibria, Redding et al. (2011) use the German division of 1949 as a natural experiment. Their research is related to our paper in that they consider the case of the central German airport (hub), which was relocated from Berlin to Frankfurt/Main in the wake of German partition and remained there even after German reunification. Redding and Sturm (2008) examine how German division and reunification affected the economic prosperity of West German cities located in the Zonal Border Area.

Becker et al (2020) caution against viewing the 1949 German division as a natural experiment in which the German population was split into two very similar subsets. Rather,

they argue, already in the interwar period, distinct differences existed between West Germany and the territory that later became the GDR. Among other things, the population in the territory of the later GDR had a lower self-employment rate, higher female employment, more Protestants, and a greater preference for left-wing parties than the population in West Germany, the later Federal Republic. This legitimate criticism does not affect our study because we focus on the location decisions of East German firms that sought their fortunes in West Germany after 1945, but not on the long-run effects of firm migration on economic development in the regions of origin and destination.

2 Historical Background

Between spring 1945 and 1950, about 18,000 companies relocated from the Soviet Occupation Zone/GDR and East Berlin to the West Zones/Federal Republic and West Berlin according to estimates by Held (1956, p. 323 f.). This westward migration took place in three main phases (Bähr 1997, p. 233). Between February and July 1945, firm owners and managers spontaneously fled the approaching Soviet army or were sometimes even evacuated to the West by the American occupation forces against their expressed wishes. Between 1946 and 1948, many East German companies reacted to dismantling, expropriation, and the gradual transition to a centrally planned economy by moving to the free-market West, where independent and profit-oriented entrepreneurial activity remained possible. The final division of Germany in 1949 destroyed any hope of the continued existence of a unified German state and therefore prompted further entrepreneurs and managers who had previously remained in the East to relocate their businesses to West Germany.

The relocation of companies to the West could take various forms (Bähr 1997, p. 231). In many cases, owners just decided to relocate their company's official headquarters to the West, which initially preserved the legal unity of the company branches located in the West and the East. This was not the case when former employees or the expropriated firm owner himself undertook a re-establishment of the company in the West, so that henceforth legally independent Western and Eastern versions of the company coexisted and sometimes even cooperated, at least in the early years of the German division. For large companies, the relocation to the West was facilitated by already existing Western branches, one of which could form the nucleus for the new headquarters. In this case, the choice of location in West Germany was not independent but affected by the existence of branches whose sites had already been selected at an earlier stage under completely different conditions. Finally, the relocation also

affected the eastern branches of West German and foreign companies. The German sewing-machine factory of the US Singer Corporation, for example, was moved from Wittenberg in the East to Karlsruhe in the West.⁷

Held (1956, p. 317) took the figure of about 18,000 companies that left the later GDR and East Berlin between 1945 and 1950 from unpublished statistics of the West German census of September 1950. By this time, a total of more than 85,000 companies had been relocated to the Federal Republic and West Berlin.⁸ The areas of origin of these relocated enterprises were, in addition to the Soviet occupation zone/GDR (11,123) and East Berlin (6,535), the territories east of the Oder-Neisse line (37,085), Czechoslovakia (18,292), the then independent Saarland (263), and other territories (including Poland, Hungary, and Austria, 11,829) (Held 1956, p. 323 f.). These figures show that it was by no means only large companies that migrated to the West. The companies that fled the Soviet occupation zone or the later GDR and East Berlin, for example, included not only well-known names such as the industrial companies *Agfa*, *Auto Union*, *Carl Zeiss* and *Knorr Bremse* or banks and insurance companies such as *Deutsche Bank*, *Dresdner Bank* and *Gothaer Versicherungen*, but also many small and medium-sized commercial firms. Hefe (1998, pp. 60, 65) estimates that about 30 to 40 percent of commercial firms left the later GDR, the vast majority of which sought a new location in West Germany rather than abroad.

The major German banks, which had little to hope for from a socialist occupying power and which, unlike many industrial companies, were not spatially tied to any physical production facilities, were particularly quick to decide to relocate to the West. For example, as early as February/March 1945, *Deutsche Bank*, based in East Berlin, moved its assets and securities holdings to Hildesheim and a replacement headquarter to Hamburg, both located in what would later become the British occupation zone (Holtfrerich 1995, pp. 420-423).

The relocation of East German companies to the West reached a first peak in the course of the withdrawal of the American Army from the highly industrialized areas of Thuringia and Saxony in the summer of 1945. By the end of the war, the Western Allies had advanced fast towards the river Elbe and occupied territories that had been earmarked for the Soviet Union according to the First London Protocol of September 12, 1944, in which the Allies had defined

⁷ See the administrative document “Bedeutung des Denkmals SINGER-Nähmaschinenfabrik AG Wittenberg“ issued by the *Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Museum* in July 2011. <https://www.veritas-park.de/images/pdf/Historie-Denkmal-Amt.pdf>, accessed on July 31, 2023.

⁸ This figure does not include those agricultural businesses or commercial enterprises that were newly founded by refugees after their arrival in the West but had no connection to an Eastern predecessor company.

the western borders of the intended Soviet-Russian occupation zone.⁹ The withdrawal of the Western troops was the subject of feverish negotiations in the weeks that followed, and rumors of the imminent entry of Soviet troops also filtered through to firms based in Saxony and Thuringia.

For example, the chairman of the board of *Vereinigte Glanzstoff AG* (one of the leading producers of artificial fibers), Ernst Hellmut Vits, learned as early as May 29 of the planned surrender of Thuringia to the Soviet Russians and immediately ordered the transfer of important machinery and documents from the Thuringian plant in Elsterberg to Oberburg/Main in Bavaria (Henke 2009, p. 732 f.). At the same time, *Siemens* began to dismantle its production sites in the Gera-Erfurt area. By truck and rail, 2,000 to 3,000 company employees with their families and 500 tons of machinery and patent files were moved from Thuringia to Bavaria (Henke 2009, p. 734 f.).

The American military administration tolerated this transfer of machines, technological know-how and human capital, which increased the economic potential of its own occupation zone and reduced that of the Soviets. If the relocation to the West did not take place on the firm owners' or managers' initiative, the Americans did not hesitate to force it even against the will of those involved. For example, the forced evacuation of the companies *Carl Zeiss* (optical instruments) and *Schott* (glass production) from Jena was organized just a few days before the official American withdrawal, which was scheduled for July 1, 1945. In addition to tons of optical and precision mechanical instruments, tens of thousands of patent specifications and design drawings, were transferred to the US occupation zone along with many managers, scientists, and skilled workers. Overall, about 1,700 people arrived in the US occupation zone in Heidenheim (Württemberg). Just fifteen kilometers north of Heidenheim, in Oberkochen, the new headquarters of the new western *Carl Zeiss* company was established in mid-1946. *Schott AG* finally found its new home in Mainz (Henke 2009, p. 760 f.).

The second phase of the relocation of East German companies began at the latest with Order 124 of the Soviet Military Administration of October 30, 1945, which set the stage for seizing companies from active National Socialists and "war profiteers." According to Steiner (2007, p. 46), almost all large enterprises and a large fraction of medium-sized businesses fell victim to confiscation and, from mid-1946, to expropriation. The so-called Soviet joint stock companies represented a special form of expropriation (Steiner 2007, p. 34 f.). By Order 167 of the Soviet Military Administration of June 5, 1946, some 200 East German industrial

⁹ See https://germanhistorydocs.ghi-dc.org/pdf/eng/Allied%20Policies%201_ENG.pdf, accessed on August 11, 2021.

enterprises originally intended for dismantling were converted into joint stock companies owned by the Soviet Union. The production of these companies was to be used to cover the reparation claims of the occupying power. One of the industrial companies affected by this development was the *Agfa* film factory in Wolfen near Bitterfeld, which was absorbed into the Soviet joint stock company *Photoplenka* in mid-1946 and only became a state-owned enterprise of the GDR in 1954 (Karlsch 1991). However, *Agfa* executives and part of the scientific and technical staff had fled to the West German city of Leverkusen, where a new film company, *Agfa AG für Photofabrikation*, was established in April 1952. On March 20, 1953, *Agfa Camera Werke AG*, a manufacturer of cameras, was founded in Munich. The two new West German companies merged on April 1, 1953. In the years that followed, the West German and East German companies fought over the use of the *Agfa* trademark. It was not until 1956 that an agreement was reached. From then on, the East German company was allowed to use the brand name in the socialist economic area and the West German company in the rest of the world.

The gradual transition to a Soviet-style centrally planned economy, which culminated in the founding of the GDR on October 7, 1949, prompted many firm owners to relocate to West Germany from 1946 onward, where it was still possible to operate a profit-oriented private company.¹⁰ In their wake, many professionals and skilled workers from East Germany followed until the government of the GDR stopped this brain drain by the closure of all remaining border crossing points and the construction of the Berlin Wall in 1961. During the second phase of relocation, it was mainly human capital that migrated. The managers, specialists, and skilled workers used their knowledge and experience to rebuild their companies in the West. Usually, the bulk of the physical capital stock, including plants, heavy equipment, and most of the machinery had to be left behind. It was either dismantled by the Soviets or continued to be used as part of the nationally owned enterprises in the East. However, in most cases, the remaining firms quickly lost their international competitiveness, which could have been either the result of the lack of human capital due to the enormous brain drain to the West or the result of the negative effects of a socialist economy.

Many other examples of relocated firms can be listed. The main plant of the renowned hosiery manufacturer *ARWA* in Auerbach (Erzgebirge) in Saxony was expropriated in the summer of 1946 and later placed under the control of the *Vereinigung Volkseigener Betriebe Trikotagen und Strümpfe*. Hans Thiersfelder, the grandson of the company founder A. Robert

¹⁰ On the reasons for leaving East Germany expressed by businessmen in personal conversations with the author, see Held (1956) p. 321. See also Dittrich (1951).

Wieland, reestablished the *ARWA* company near the city of Gaildorf in Württemberg in 1949.¹¹ In September 1949, together with core workers who had also migrated, the former Chairman of the Board of Management of *Chemnitz Auto Union AG*, Richard Bruhn, and his deputy, Carl Hahn, founded *Auto Union AG* in Ingolstadt, Bavaria, which would eventually become the headquarters of today's *Audi AG* (Mirsching 1988). In 1946, the famous doll maker Käthe Kruse, whose production had until then taken place in Bad Kösen (Saxony-Anhalt), sent her two sons Max and Michael to the western occupation zones to set up branch factories there. In 1950, the entire production of Käthe Kruse dolls was transferred to the branch factory opened by Michael Kruse in Donauwörth (Bavaria); the East German parent company was expropriated in 1952.¹² In 1872, *Baedeker Verlag*, a publishing house specialized in travel literature, which was originally founded in Koblenz (Rhineland), had moved to Leipzig, the traditional center for book trade in Germany. However, the great-grandson of the company's founder, Karl Friedrich Baedeker, saw no future for the publishing house in the Soviet occupation zone and moved the company's headquarters again in 1948, first to Malente in Schleswig-Holstein, and then in 1956 to Freiburg in Baden.¹³ Another example is the machine factory *Max Loesch* from Dresden, which had specialized in the manufacture of packaging machines for the confectionery industry since 1919. In 1949, two of its employees, Erhardt Walther and Erich Adler, reinstated the company in Upper Franconia after it had been expropriated by the Soviets.¹⁴ After the confiscation of the Magdeburg plant of the venerable *Buckau R. Wolf* machine factory, its director Wilhelm Kleinherne moved to the western branch in Grevenbroich, which became the new official headquarters of the company in 1947 (*Maschinenfabrik Buckau R. Wolf* 1963). The Chemnitz machine tool manufacturer *Reinecker AG* was severely damaged during World War II and the remaining facilities were dismantled in the post-war period on the orders of the Soviet military administration. In 1949, former CEO Rudolf Westenberger and the husband of the founder's granddaughter, Curt Kloetzer, rebuilt the company in Munich as *J. E. Reinecker Maschinenbau GmbH*. Facilities to produce milling machines were established in Ulm-Einsingen.¹⁵

This list could be continued for many hundreds of companies. Our examples can only give a first idea about the wide range of firms from different industries and size classes that

¹¹ [https://saebi.isgv.de/biografie/Hans_Thierfelder_\(1913-1987\)](https://saebi.isgv.de/biografie/Hans_Thierfelder_(1913-1987)), accessed on August 11, 2021.

¹² <https://www.kaethe-kruse.de/de/geschichte#>, accessed on August 11, 2021.

¹³ <https://www.baedeker.com/verlag/>, accessed on August 11, 2021.

¹⁴ <https://www.loeschpack.com/en/company/history.html>, accessed on August 11, 2021.

¹⁵ <https://daten.digital-sammlungen.de/0001/bsb00016339/images/index.html?seite=363>, accessed on August 11, 2021.

migrated to the West. In the following, we describe our data set on relocated East German firms in more detail.

3 Data on Relocated East German Firms

To analyze the choice of relocation of East German firms after World War II, we construct a novel data set that is based on the “Register of west addresses of industrial and trade companies, banks, and insurance companies from the Soviet-occupied zone and the East sector of Berlin”. This register was published in 1957 by the business association of companies that had been expropriated in Eastern Germany.¹⁶ It lists over 5,200 relocated firms with information about the place where the firm was based in East Germany, the new firm address in West Germany (and the new company name in case that it changed), the related industry, and whether the firm was still in operation when the register was published.¹⁷ Based on this data, we assign geo-codes to each firm’s old location in East Germany and its new locations in West Germany. In addition, we standardize information about the firms’ related industry by assigning a Standard Industry Classification (SIC) code to each firm. We use the company name to identify whether the firm was a joint-stock company (AG), a limited liability company (GmbH), or a privately-owned company.

Table 1 gives an overview of the data. We differentiate between all firms included in the register and the subset of firms that were still in operation in 1957. We also distinguish between firms that were relocated to West Germany (excluding West Berlin) and firms that were relocated to West Berlin. Motives to relocate to West Berlin might have been different because this city was an exclave surrounded by hostile socialist territory. Overall, we observe the new locations of 4,676 East German firms in West Germany with 3,765 that were still in business in 1957, and 506 East German companies that moved to West Berlin with 449 that were still in business in 1957. Table 1 also shows the distribution of firms across company types (AG, GmbH, and privately-owned firms). Most of the firms in the sample were privately-owned, while AGs and GmbHs only account for a small fraction. However, when compared with the distribution of all firms that existed in inter-war Germany, AGs and GmbHs are clearly overrepresented in our sample. On average, AGs and GmbHs are larger than private-owned

¹⁶ See Interessengemeinschaft der in der Ostzone Enteigneten Betriebe (1957). *Westadressen-Verzeichnis der Industrie- und Handelsbetriebe, Banken und Versicherungen aus der sowjetisch besetzten Zone und dem Ostsektor Berlins*. Verlag Klaus Edgar Herfurth, Frankfurt/Main.

¹⁷ For additional information, see Appendix B1. To illustrate the structure of the original data source, we show the first ten firm entries, as reported in the register of West addresses, in Table B1 in the appendix.

firms.¹⁸ Larger firms might have been more inclined to relocate their business to the West which would explain the exceptionally large number of AGs and GmbHs in our data.¹⁹

[Insert Table 1 here]

Figure 1 shows the distribution of relocated companies across industry groups, which we aggregated following SIC (see Table B2 in the Appendix for more details). Overall, the manufacturing sector accounts for the largest fraction of all firms with a particularly large number of textile and machine-building firms, but the sample also includes a considerable share of wholesale and retail trading firms as well as other service-sector firms. The number of relocated agricultural firms is small, which reflects the fact that land is not mobile. Given that agriculture was still a quantitatively relevant sector in East Germany, our sample of firms is obviously not representative regarding the relative importance of agriculture. The distribution of relocated firms across industries in manufacturing, however, seems to resemble the overall distribution of manufacturing firms in pre-war East Germany quite well.²⁰

[Insert Figure 1 here]

In Figure 2 and 3, we provide information about the spatial distribution of relocated firms. Figure 2 shows the places to which East German firms migrated after the war; Figure 3 displays the respective places of origin in pre-war East Germany. In Figure 2, the spatial distribution resembles the economic geography of West Germany with a high number of firms relocated close to the gravity centers of economic activity such as the Rhine-Ruhr region, the area around Frankfurt am Main, and the most industrialized parts of Southwest Germany. However, we also observe a considerable number of firms that migrated to places close to the East-German border, even though there existed no major agglomerations in these regions. This later finding suggests that geographic proximity to the pre-war locations might have also played a role in East German firms' relocation choice. Figure 3 shows that many firms came from Saxony and Thuringia, which had been highly industrialized region since the 19th century. By contrast, only few firms were from the northern part of East Germany (Mecklenburg, Pomerania, and parts of Brandenburg), where the economy was still dominated by agriculture.

¹⁸ In 1925, privately owned firms (sole proprietors and private partnerships) accounted for 96.3% of all firms but only for 59.1% of all employees. The share of GmbHs (AGs and KGaAs) was 1.3% (0.4%) but they accounted for 8.6% (20.3%) of all employees. Data source: Statistisches Bundesamt (1954).

¹⁹ In Appendix A3, we investigate the influence of firm type on location choice.

²⁰ Ritschl and Vonyó (2014, p. 177) show that, in 1936, the industries textiles, machine-building, foodstuff, drinks & tobacco, printing & publishing, and chemicals accounted for an above-average share of gross value added in East Germany.

[Insert Figures 2 and 3 here]

We match each place of a relocated firm to the respective county in West Germany based on the administrative structure of the early 1950s. In total, there are 467 counties, thereof 395 counties for which we observe at least one relocated firm and 72 counties with no relocated firm. Table 2 shows descriptive statistics for the number of firms by county, including information for different subsamples of firms (AG, GmbH, privately-owned firms, manufacturing firms, and non-manufacturing firms). We include only firms that were still in operation in 1957. The mean number of firms per county is 9.12 for the full sample (all firms).

[Insert Table 2 here]

4 Description of the Model

We assume that rational decision-makers, whether the firm owners or managers, compare the expected profits that their firm can generate at each potential location. Based on this comparison, they select the location that maximizes the expected profits. In our specific setting, this means that the decision-maker of firm i , located in a specific East German place, chooses a new firm location in a West German county c out of a set of 467 alternative counties ($c = 1, 2, \dots, 467$). In our baseline model, we assume for simplification that the expected profits that firm i generates in county c depend on two parameters: the distance between the original home location of firm i and county c (d_{ic}) and the size of the agglomeration of county c (a_c). The expected profit π_{ic}^e of firm i in county c is then given by:

$$(1) \quad \pi_{ic}^e = f(d_{ic}, a_c)$$

In locations geographically close to their home regions, firms may benefit from existing networks with customers and suppliers and other information advantages which lower their transaction costs. Thus, we assume that d_{ic} affects expected profits negatively so that the first derivative of π_{ic}^e with respect to d_{ic} is negative:

$$(2) \quad \frac{\partial \pi_{ic}^e}{\partial d_{ic}} < 0.$$

In large agglomerations, firms gain among other factors from a better supply of skilled labor and intermediate products, from more customers, from a better infrastructure, and knowledge spillovers. We assume that profits increase with the size of the agglomeration so that the first derivative of π_{ic}^e with respect to a_c is positive:

$$(3) \quad \frac{\partial \pi_{ic}^e}{\partial a_c} > 0.$$

By using the parameter δ for the profit elasticity of distance (with $\delta < 0$) and α for the profit elasticity of the agglomeration advantages (with $\alpha > 0$), the profit equation is given by

$$(4) \quad \pi_{ic}^e = d_{ic}^{\delta} a_c^{\alpha},$$

which we transform by taking the logarithm on both sides:

$$(5) \quad \ln(\pi_{ic}^e) = \delta \ln(d_{ic}) + \alpha \ln(a_c)$$

The decision-maker chooses the firm location by selecting the county with the highest expected profits. $Choice_{ic}$ is the variable that indicates the choice of firm i . It equals 1 for the county that maximizes the expected profits and 0 for all other counties:

$$(6) \quad Choice_{ic} = \begin{cases} 1, & \ln(\pi_{ic}^e) = \max[\delta \ln(d_{ic}) + \alpha \ln(a_c)] \\ 0, & \ln(\pi_{ic}^e) < \max[\delta \ln(d_{ic}) + \alpha \ln(a_c)] \end{cases}$$

Note that the parameters δ and α differ across industries. For simplification, we distinguish between two types of industries, *LOCAL* (L) and *GLOBAL* (G). In industry L , firms produce mainly for local markets (e.g., the manufacturer of consumer goods for which strong regional preferences exist) or rely on a local supplier base (e.g., wood-processing firms that buy wood from local suppliers), while, in industry G , the firms produce mainly for export markets (e.g., chemical firms) or rely on an international supplier base (e.g., automotive firms that buy components from suppliers from all over the world). Thus, we assume that $|\delta_L| > |\delta_G|$, which implies that distance has a stronger negative effect on expected profits in industry L than in industry G . Firms in industry L also differ from firms in industry G with respect to the expected positive effects of agglomerations. We assume that $|\alpha_L| < |\alpha_G|$, which means that the positive agglomeration externalities are smaller for firms in industry L compared to firms in industry G (e.g., because of a lower demand for highly qualified labor, which is available in large quantities only in agglomerations).

Since $|\delta_L| > |\delta_G|$ and $|\alpha_L| < |\alpha_G|$, firms in industry L choose locations close to their original place, for which d_{ic} is sufficiently small, with a higher probability when compared to firms of industry G , for which the “penalty” of distance is weaker, while the positive agglomeration externalities are stronger.

To estimate the coefficients with our data, we use a mixed multinomial logit model (McFadden and Train 2000). Based on the underlying profit function, the model that we estimate takes the following form:

$$(7) \quad \text{Choice}_{ic} = \delta D_{ic} + A'_c \alpha + X'_c \beta + \varepsilon_{ic}$$

Choice_{ic} is a dummy variable that equals one for county c that firm i chooses, and zero for all other observations. D_{ic} is the distance between county c and the old location of firm i , measuring potential home advantages. A'_c includes different measures for the potential agglomeration advantages of county c . In addition, we include a vector with further determinants of location choice (X'_c) that vary across counties to measure effects that are not captured by our agglomeration proxies. ε_{ic} is a random term, δ is the coefficient for the effect of distance, α is a vector with coefficients for different agglomeration proxies, and β is a vector containing the coefficients for all other controls. In the main sample, we observe 4,214 firms. Since each of these firms can choose among 467 alternatives, the model includes 1,967,938 observations.

5 Determinants of Firm Location Choice

5.1 Distance as Proxy for Home Advantage

One of our main variables of interest is the potential home advantage, which we measure with distance (D_{ic}). By using geo-coordinates of old firm locations in East Germany and the centroids of each county in West German, we compute the great-circle distance (in km) between the old location and each potential destination county c in West Germany for each firm i , respectively.

5.2 Agglomeration Advantages

To account for agglomeration advantages, we use three variables:

First, we take the population density in a county (inhabitants per km²). Areas with high population densities have typically many advantages including easy access to (highly qualified) employees, customers, suppliers, or a superior transport infrastructure. Since, after the war, the relocation of East German companies may have affected the population in the receiving counties, we use pre-war population data from the 1933 census to compute population density at the county level. Another advantage of the 1933 census is that, in this year, the population distribution is unbiased by war-related distortions such as the bombing war, which forced people to leave the destroyed cities for some time. Better than the 1950 census, the 1933 census reflects the persisting long-run agglomeration patterns in Germany.²¹

²¹ Another argument for the use of the 1933 census is that East German decision-makers had probably good knowledge about the broad agglomeration patterns in West Germany but lacked precise information about the changes in the agglomeration pattern between 1933 and 1950.

Second, we add a more specific proxy for local economic activity to deal with the concern that population density may not fully reflect regional differences in economic activity. For example, if the population density in county A was the same as in county B but the companies located in A were more productive (e.g., due to a better-educated workforce), a firm may have chosen county A over B. Since county-level GDP data is not available for the pre-war period, we compute firm revenues per capita from the 1935 revenue tax statistic to control for differences in regional productivity not captured by population density. Again, we rely on pre-war data to avoid concerns about endogeneity, because the relocation of East German firms after 1945 clearly affected local firm revenues in the West.

Third, it is likely that the decision to relocate a firm was not only influenced by the agglomeration advantages of a specific county but also by the characteristics of the surrounding counties. For example, if the population density in county A was the same as in county B but the surrounding regions of A were highly populated while B was the middle of a sparsely populated agricultural area, a firm may have chosen county A over B because A promised both more regionally available workers and more easily accessible customers. To control for this effect, we use the variable Domestic Market Potential (DMP), which we define in the following way:

$$(8) \quad DMP_c = \sum_{j \neq c} \frac{Pop_j^{1933}}{(D_{jc})^2}$$

Pop_j^{1933} is the county population of county j , which we divide with the squared distance between county j and c . We use the squared distance in the denominator to give more distant counties a lower weight. We compute the DMP for each county c by summing up all distance-weighted county populations.

While we prefer data from the 1933 census to avoid endogeneity concerns, we computed all three agglomeration variables also for 1950, using data from the West German census, to test the robustness of our results.²²

5.3 Other Determinants

The decision to relocate a firm to a specific county might have also been driven by other factors than the home advantage, measured by the distance to the old place, or agglomeration effects. To capture other potential determinants, we add a battery of additional controls.

²² In Appendix A1, we show that the data are highly correlated (when comparing 1933 and 1950, see Figure A1 in Appendix A1) and that the main results are similar, when using data from 1950 for the regressions. In doing so, we also control explicitly for the change in the respective variables between 1933 and 1950 (see Table A2 in Appendix A1).

First, we must consider institutional differences between the occupation zone, which may have affected the decision to relocate a company. We add dummy variables for Berlin, the British occupation zone, and the French occupation zone (with the US occupation zone as reference group). By controlling for the French occupation zone, we account, for example, for potential effects that arose from the ban on immigration, which is prominently discussed in the literature (e.g., Schuman 2014, Ciccone and Nimczik, 2022).

Second, since differences in the local availability of human capital may not yet be fully captured, we control for potential access to university graduates by using alternatively a variable that indicates whether a university was in the respective county in 1950 or a variable that measures the minimum distance between each county and a county with a university in 1950.

Third, East German decision-makers may have preferred counties with superior transport infrastructure. While differences in population density should largely reflect differences in traditional transport infrastructure, since railways or access to waterways affected population growth positively (see, e.g., Donaldson and Hornbeck, 2016; Hornung, 2015), we use additional dummy variables that indicate major seaports and access to motor highways, respectively. From the perspective of the late 1940s, access to motor highways, which were constructed before the war, may have been especially relevant in the decision process since these highways were an innovative mode of transport. Given that the highways were constructed in the 1930s, the potential advantages of highway access are not yet captured by the population density in 1933.

Fourth, we include the latitude and the longitude of the respective county as controls. By controlling for latitude (longitude), we account for north-south (west-east) differences that are not reflected in other variables. Counties in the west of Western Germany, for example, may have profited from geographical proximity to France, Belgium, and the Netherlands, which may have facilitated trade with these countries. Better access to other Western European markets may have in turn influenced the location choice of an East German decision-maker.

Fifth, financial incentives may have affected the location decision. Since West German politicians were worried that areas close to the new established West-East-German border could suffer economically, they introduced a specific subsidy program for firms in this area, the so-called “zonal area funding” (*Zonenrandgebietförderung*). The West German government decided for subsidies of 100 million DM in 1951 and the parliament approved the funding in 1953. From the perspective of the late 1940s, decision-makers may have already expected such kind of subsidies in their decision process. Therefore, we use a dummy variable indicating all counties at the East border that were eligible for border area funding.

Apart from subsidies, differences in business taxation may have affected the location decision (Devereux and Griffith, 1998). To account for this effect, we exploit county-level differences in the German business tax rate (*Gewerbesteuer*). Local business tax rates were set by municipalities, which attempted to attract new businesses by setting comparatively low business tax rates. For this reason, we also understand this variable as a measure of all other unobservable industrial policy measures that local administration took to increase the attractiveness of their local location.

Last, we test whether the decision to relocate a firm to a specific county was affected by the share of expellees in the respective county. Previous research has studied the forced migration of Germans that had to leave the Eastern parts of pre-war Germany (including the area East of the Oder-Neisse line and the Sudetenland) and settled down in the West (e.g. Braun and Franke 2021, Braun et al. 2021). Counties that received many expellees may have profited from the inflow of people, increasing the available workforce and human capital in the respective counties.

6 Empirical Results

To identify the determinants of firm location choice, we apply the mixed multinomial logit model that we have described in section 4. In doing so, we estimate the likelihood with which an East German firm chose a specific county in West Germany based on the determinants of firm location choice that we have discussed above. In Table 3, we show descriptive statistics for all determinants that we include in our regressions. To facilitate the interpretation of the coefficients, we take the natural logarithm of all continuous variables. The number of observations (1,967,938) is equal to the number of relocated firms (4,214) multiplied by the number of choices (467 counties). For some variables, we have no data for specific counties so that the number of observations decreases accordingly.

[Insert Table 3 here]

6.1 Main Results

Table 4 shows the estimated coefficients for the main model. In the first column, we report the estimates for the basic model specification including only distance as proxy for the home advantage and the population density in 1933 to measure agglomeration effects. The results show that the effect of $\ln(\text{Distance})$ is significantly negative, suggesting that proximity to the original East German home market was crucial in the decision where to relocate the firm. The coefficient for $\ln(\text{Pop. Density 1933})$ is significantly positive, as suggested by the literature on

the economics of agglomeration. In column (2), we add $\ln(\text{Revenue pc } 1935)$ as an additional control. We use this variable as proxy for GDP per capita, indicating differences in firms' productivity between counties. The estimated coefficient is significantly positive. While the estimated coefficient of $\ln(\text{Distance})$ remains almost unchanged, the magnitude of the agglomeration effect, $\ln(\text{Pop. Density } 1933)$, decreased by about 50 percent, but it remains significant. In column (3), we also include $\ln(\text{DMP } 1933)$, which measures the Domestic Market Potential based on the distance-weighted population in surrounding counties. When controlling for the Domestic Market Potential, the negative effect of $\ln(\text{Distance})$ increases.

In columns (4) to (8), we include additional variables that may have also affected the location choice. First, in column (4), we account for differences in the occupation zones after World War II by adding the dummy variables *Berlin*, *British Occupation*, and *French Occupation*. The coefficients indicate that, when controlling for all other factors, it was less likely that a firm chose the British occupation zone, compared to the US occupation zone, which is the reference category. The negative effect of *French Occupation* is not significant. The positive effect of *Berlin* is relatively strong and significant, which can be explained by the fact that many companies from the formerly economically strong Greater Berlin area were looking for a nearby non-socialist haven in West Berlin. Interestingly, the inclusion of dummies for the occupation zones leads to remarkable changes in the coefficients of the variables measuring the impact of agglomeration effects. The coefficient of Domestic Market Potential quadruples while the coefficient of Population Density becomes significantly negative. One possible explanation for this surprising result is that by introducing other variables that more accurately capture the economic effects of agglomeration, population density primarily measures the negative effects of overcrowding.

In column (5), we add proxies for access to the knowledge available in nearby universities and to the graduates trained there by using the dummy variable *Uni 1950*, which has a significantly positive effect and $\ln(\text{Distance to Uni } 1950)$, which has no significant effect. This observation supports the notion that the presence of a university makes counties more attractive to businesses. In column (6), we also include the dummy variables *Sea Port* and *Highway*, which both affect the location choice significantly positive. Next, in column (7), we include *Latitude* and *Longitude* as controls. The effects of latitude and longitude are significantly negative, suggesting that, after controlling for all other factors, firms preferred destinations in the West and South of Germany. Last, in column (8), we add the dummy variable *Eastern Border*, which indicates counties that were located directly at the border to the later GDR. The effect of *Eastern Border* is significantly positive. One explanation for this effect

could be that decision-makers were attracted by the expected subsidies from the Zonal Area Funding Program. In all five columns, (4) to (8), the effect of $\ln(\text{Distance})$ remains significantly negative and the coefficient size is hardly affected by the inclusion of additional controls. In contrast, the magnitudes of the coefficients of the three agglomeration proxies decrease in comparison with column (3) but remain significant.²³

[Insert Table 4 here]

In Table 5, we include further controls that are not available for the full sample. First, we explore the effect of local taxes by exploiting county-level variation in the German business tax rate.²⁴ Data on the local business tax rate is only available for 283 counties. To enable an appropriate comparison, we first estimate the model with all previous controls from column (8) of Table 4 for the restricted sample. Column (1) of Table 5 reports the results for the coefficients of $\ln(\text{Distance})$ and the three agglomeration proxies. In column (2), we then add the variable *Tax Rate*, which has a significantly negative effect on the location decision, while the effects of $\ln(\text{Distance})$ and the agglomeration measures remain unchanged. Second, we test whether the share of expellees from former East German territories is correlated with the location choice. Since expellee data is not available for West Berlin, we first estimate the model with full controls for the restricted sample that excludes Berlin in column (3) of Table 5. When dropping West Berlin from the sample, the absolute value of the effect of $\ln(\text{Distance})$ increases from 1.191 (column (8) of Table 4) to 1.440. In column (4), we then include *Expellees 1950* as an additional control. The effect of *Expellees 1950* is significantly positive. The effect of $\ln(\text{Distance})$ remains almost unchanged compared to column (3) of Table 5, while the negative effect of population density becomes insignificant.

[Insert Table 5 here]

The previous tests have shown that the coefficients of the main variables of interest are quite stable and hardly change when we include additional variables or vary the sample size. In addition, we show in Appendix A1 that we find similar results when computing all agglomeration proxies based on post-war population and firm-revenue data.

²³ Note that the effect of *French Occupation* gets significantly negative, when adding the controls in column (7) and (8). This result may reflect that the French occupation zone remained closed to refugees until 1949, as noted by Held (1959).

²⁴ The business tax rate differed in Germany because every community could define a “tax factor” that increased the business tax payable (the so-called *Hebesatz*). See Appendix B5.7 for more details on the German business tax.

To test the predictive power of our model, we use the estimated parameters from column (8) of Table 4 to compute the share of firms that choose county c relative to the total number of relocated firms for each county, respectively. We then compare the predicted shares with the actual (historical) shares that we can observe from the data. Figure 4 plots the results with the natural logarithms of the actual shares on the y-axis and the predicted shares on the x-axis.²⁵ When we regress the natural logarithms of the predicted shares on the actual shares, we get a coefficient of 1.0449 and an R^2 of 0.98, showing that the model predicts the actual variation quite well.²⁶

[Insert Figure 4 here]

Throughout all specifications, we find a significantly negative effect of distance on the location choice, suggesting that the home advantages matter, while the effects for the agglomeration proxies are mixed. To evaluate the economic magnitude of the estimated coefficients of our main variables of interest, we first transform the coefficients of the full-sample model with all controls in column (8) of Table 4 into odds ratios (OR) and then multiply the odds ratio with one standard deviation. Table 6 shows the results of this calculation for $\ln(\text{Distance})$ and all three agglomeration proxies.

[Insert Table 6 here]

For the estimated effect of $\ln(\text{Distance})$, the odds ratio is 0.30. If we increase $\ln(\text{distance})$ by one standard deviation (0.42), then the odds of choosing a corresponding county would decrease by $0.42 \times (0.3 - 1) \times 100\% = -29.4\%$. To illustrate the magnitude of this effect, we take a firm that was originally located in the East German city of Leipzig and compare two potential choices whose logarithmic distance to Leipzig differs by just one standard deviation. These are, for example, Bayreuth (in northern Bavaria), which is closer to Leipzig, and Detmold (in Westphalia), which is farer away. Thus, on average, when holding all other factors constant, the odds that the firm from Leipzig chooses Detmold over Bayreuth are 29.4 percent lower according to our model. This effect is economically large, suggesting that proximity to the former home market was an important determinant in the decision to relocate the company.

²⁵ We take the logarithm to give counties with a high share a lower weight and to facilitate the presentation of the data. Before taking the logarithm, we have added 0.0001 to the actual and predicted shares, respectively, since there are counties for which the actual share is zero.

²⁶ If we run a similar univariate regression without taking the natural logarithms of the respective shares, the estimated $\hat{\beta}$ is 1.0176 (standard error: 0.0411) and the R^2 is 0.93, showing that the result is not driven by the log transformation.

Similarly, we can compute the economic magnitudes of the agglomeration effects. The effect of $\ln(\text{Pop. Density } 1933)$ is negative, but the coefficient is small so that a one-standard-deviation increase lowers the odds to choose a county by only 7.3 percent. By contrast, the positive effect of $\ln(\text{Revenues pc } 1935)$ is economically large. An increase by one standard deviation increases the odds to choose a county by about 75 percent. The effect of $\ln(\text{DMP } 1933)$ is of similar size. Thus, the negative effect of the population density—which seems to be surprising at first sight—is far too small to compensate for the other positive effects of agglomeration that occurred in populous counties that were characterized by high productivity of local businesses or by being adjacent to other populous counties.

To conclude, under otherwise identical conditions, East German firms' decision-makers preferred West German counties that were close to their former home markets to counties that were far away, and they preferred counties that were part of a larger regional agglomeration with highly productive firms to populous counties with less productive firms and with a location at the periphery. In comparison, the home advantage effect appears to be weaker than the agglomeration effects. However, we must consider that it was geographically impossible for an East German company from the “far” east of what would later become East Germany to select a nearby county in West Germany. This limitation weakens the negative effect of distance. In the following two sections, we will examine the apparent non-linearity of home advantages in more detail.

6.2 Non-Linear Distance Effect

In all previous specifications, we have used the natural logarithm of the distance between an East German firm's old location and its potential new location to estimate the effect of distance. To test for potential non-linear distance effects, we now use a set of dummy variables indicating whether the distance between the old place and the potential new place was in a specific 50 km interval. The first dummy variable covers observations for which the distance is up to 50 km, the next interval includes distances between 50 and 100 km, and the last interval indicates distances above 600 km. The 300 to 400 km interval serves as the reference category, which includes the median distance (about 342 km). We expect the coefficient of the distance dummies to be positive for very short distances because companies preferred locations close to their old home base. With increasing distance, the effect should become smaller and eventually negative for large distances. To test for this hypothesis, we rely on the baseline specification with all controls (see column (8) of Table 4) but include the distance-interval variables instead of $\ln(\text{Distance})$. In Figure 5, we present the estimated coefficients for eleven distance intervals.

[Insert Figure 5 here]

Figure 5 shows that the effect of distance gets indeed smaller when we increase the distance. For a better interpretation of the estimated coefficients, we use again odds ratios. For the first dummy variable, indicating distances up to 50 km, the estimated coefficient is about 3.8, which equals an odds ratio of 44.3, meaning that the odds that an East German firm chose a new location within 50 km to its old location was about 44 times higher than choosing a location from the reference interval. For distances between 50 and 100 km, an odds ratio of about 7.3 implies that the odds that a county from this interval was chosen were still more than seven times higher than selecting a county from the reference interval. For large distances, we find a negative effect of distance. For distances larger than 600 km, for example, we estimate a coefficient of about -1.1, which translates to an odds ratio of 0.34, meaning that the odds that a firm has chosen a county more than 600 km away from its old location were 66 percent lower than choosing a county that was in the 300 to 400 km range.

6.3 Effect of Pre-War Firm Location

If the home advantage was a crucial factor in an East German firms' location choice, we would assume that the effect of distance is stronger for firms that were originally located in areas close to the West-East border that was established in 1945. Because many of these companies may have traditionally had many customers and suppliers in nearby areas that became part of West Germany after 1945, the decision to locate just across the West-East border was an obvious one. In contrast, firms that were originally located far east of the West-East border and therefore had fewer economic contacts into the later West German zonal border area may have felt less incentives to relocate to Western areas near this border. Thus, the negative effect of distance should be weaker for the latter. To investigate whether the size of the distance effect depends on the original location of the East German firm, we use a sample-split test that distinguishes between firms that were originally in areas close to the border and those that were not.

For this purpose, first, we subdivided the area of East Germany into grid cells of 0.5 x 0.5 degrees. Next, we assigned each geo-coded old firm location to the respective grid cells. We distinguish two samples: One sample only includes firms in grid cells at the West-East border, the second sample contains the firms of all other grid cells. The border grid cells touch the West-East border, going down from the western part of Mecklenburg at the Baltic Sea over the western parts of Sachsen-Anhalt south to Thuringia, and in west-east direction from

Thuringia to Saxony. When using a narrow grid-cell definition, with only one grid cell next to the border, we observe 423 firms in border grid cells, 3,791 firms in the rest of East Germany.²⁷

In another model, we distinguish between territories that had been temporarily occupied by the Western Allies in 1945 and territories that were either occupied by the Soviets or not occupied at the time of the armistice (see section 2 for the historical details). The territory that was temporarily under Western control overlaps with the border grid cells but spans a larger area, reaching partially to the river Elbe or even beyond. In total, there were 2,290 firms in this area, compared to 1,924 in areas that were not temporarily occupied by the Western Allies.

In Table 7, we show the results for the effects of $\ln(\text{Distance})$ and all agglomeration proxies for both sample split tests, respectively. All other controls are included but we do not report the coefficients. For comparison, we first report the estimates for the full sample in the first column. In columns (2) and (3), we then split the sample based on the grid-cell definition. In the border-grid-cell sample, in column (2), the estimated coefficient of $\ln(\text{Distance})$ is -2.36 compared to only -1.043 in the sample that includes all other grid cells. This result is in line with the assumption that the home advantage is stronger for firms in areas close to the border. By contrast, for firms from places not at the border, the negative effect of pre-war population density is smaller, while the positive effect of firm revenues per capita is larger, suggesting that these firms preferred agglomerations more than the firms in the border sample.

A striking example for the decision to relocate firms just across the border provides the East German city of Sonneberg, located directly at the inner-German border. Of the 27 firms that migrated from Sonneberg to West Germany after 1945 (and were in business in 1957), 48 percent settled in the West German city (and surrounding county) of Coburg, which was only 20 kilometres away from Sonneberg.

We find similar results in columns (4) and (5), where we split the sample according to the temporary occupation of Western troops in 1945. However, when comparing the effects of $\ln(\text{Distance})$, the difference between the two coefficients is less pronounced than in the first sample-split tests. This result is not surprising. The temporary Western occupation reached into areas far away from the later West-East border, whose firms placed less emphasis on home advantages when deciding on a location than companies located close to the border. In other words, the differences in geographic locations are much more pronounced in the first sample split test (models 2 and 3 of Table 7) than in the second (models 4 and 5 of Table 7).

²⁷ See Appendix A2 for results with a wider grid-cell definition. When increasing the border-grid-cell area, the effect of $\ln(\text{Distance})$ decreases in the border sample, which reflects the results of the section 6.2.

Both sample-split tests show that East German firms from places close to the West-East border were more inclined to migrate to nearby places in the West than firms from areas that were originally located more far in the East. This result reflects the findings of the previous section, in which we show that the effect of distance is not linear but stronger for smaller distances.

[Insert Table 7 here]

Not only the distance to the inner-German border but also the economic conditions of the original location may have played a role in the location decision of the East German companies. Especially, firms that had already benefited from agglomeration advantages at their old location might have tended to search for similar places in West Germany. To test for this effect, we split the sample in firms that came from large East German cities with more than 100,000 inhabitants, based on the pre-war population, and firms that were from all other places. Columns (6) and (7) of Table 7 show the results for this additional sample-split test. In column (6), which only includes firms from large cities, the absolute value of the effect of $\ln(Distance)$ is smaller when compared with the baseline model in column (1) but the effect of population density is not negative anymore and the positive effect of firm revenues per capita has increased. By comparison, we find the opposite for the sample with all other firms in column (7). We therefore conclude that East German firms from large cities had a stronger preference for large agglomerations in West Germany than East German firms that were originally located in smaller cities or at the countryside.

6.4 Differences across Industries

Depending on their industry, East German firms might have pursued different goals with their choice of location. For a firm in an export-oriented high-tech industry, for example, agglomeration advantages could be more important than any home advantages so that the firm's decision-maker may choose to settle in a far away West German agglomeration. By contrast, for a small trading firm, which relies on established ties to local customers and suppliers, home advantages could be larger than potential agglomeration advantages, so that the firm's decision-maker would choose a place close to the old location. To test to what extent these differences existed across East German industries, we create subsamples including only firms of a specific industry and then estimate our baseline mixed logit choice model for each subsample respectively. Figure 6 shows the results for the distance and agglomeration effects. In the upper-left panel, we report the coefficients for the effect of $\ln(Distance)$ for ten manufacturing industries as well as the total average and the average across manufacturing industries for

comparison. The other panels show the coefficients for the agglomeration effects, $\ln(\text{Pop. Density } 1933)$, $\ln(\text{Revenue pc } 1935)$, and $\ln(\text{DMP } 1933)$, respectively. We order the industries in all panels according to the size of the point estimate for the effect of $\ln(\text{Distance})$.

[Insert Figure 6 here]

The effect of $\ln(\text{Distance})$ is significantly negative in all industries but the point estimates differ remarkably. When comparing electrical-engineering firms with wood-processing firms (including, for example, the production of furniture), the point estimates range from -0.7451 to -1.7055. By and large, our results are consistent with the idea of home advantages being less important for export-oriented manufacturing firms with a “global” supplier base, as it was the case for electrical engineering or the production of scientific instruments (including precision mechanics as well as medical and optical goods), than for industries that mainly produce for domestic markets and have a more “localised” supplier base, as it was the case for the wood-processing and foodstuff industries. This distinction is confirmed from another perspective: For the wood-processing and foodstuff industries, the effect of $\ln(\text{Revenue pc } 1935)$ is much smaller when compared with the average (see the lower left panel of Figure 6), suggesting again that agglomeration advantages were in these cases less important than the advantages that resulted from doing business on a well-known home market.

In Figure 7, we show the results for six non-manufacturing industries. The effect of $\ln(\text{Distance})$ is relatively strong for transportation, which includes mainly small haulage firms, and weak for financial services, which includes banks and insurance companies. Overall, the effect of distance is stronger for non-manufacturing compared to manufacturing firms, suggesting a stronger home advantage. Concerning the agglomeration effects, we find, in most cases, coefficients of similar size. An exception is transportation, for which the effect of population density is much stronger and, thus, significantly positive, and the effect of Domestic Market Potential is significantly negative (in contrast to all other industries where this effect is positive). These observations suggest that the smaller haulage firms in our sample preferred densely populated cities in economically isolated locations to larger agglomerations—perhaps because there was less competition with large haulers in such cities.

[Insert Figure 7 here]

6.5 Similar Industry Structures

There is the possibility that firms will be attracted not to agglomerations per se, but to the many other firms from their own industry that may be located there. Indeed, it is the other firms in

the same industry that ensure that suitable skilled workers, industry-specific services, or knowledge externalities can be expected within the agglomeration (Ellison et al. 2010). In our empirical setting, this form of “co-agglomeration” could lead to an overestimation of the effect of distance because East German firms may have chosen a nearby West German location not because of its proximity but because of its similar industrial structure. To rule out this alternative explanation for the home-advantage effect, we run separate regressions for firms in specific industries, as in the previous section, but now additionally control for the relative importance each industry had in the receiving county by using data from the pre-war employment census. More precisely, we include the variable *Industry Share %* as an additional control, which is defined as the number of employees of a specific industry in a specific county in percent of the total number of employees in this county. When testing the model for the chemical industry, for example, *Industry Share %* indicates the percentage share of employees in the chemical industry in a specific county. In Figure 8, we show the estimated coefficients for the effects of $\ln(Distance)$ and our three agglomeration proxies for nine different industries, respectively. Figure 9 supplements the corresponding estimates for the variable *Industry Share %*. Note that the industry definitions differs slightly from the previous section since employment data is not available for all industries.²⁸ The underlying models include all previous controls (as in column (8) of Table 4) as well as *Industry Share %*.

Figure 8 provides similar results to Figure 7 which implies that the industry structure of a county, measured by the respective pre-war employment share, does not drive our results. As shown in the previous section, for example, the home-advantage effect is particularly strong in the wood-products industry, while it is relatively weak (though still significant) for firms in the electrical engineering and instrument industries, which were more world-market oriented than the former. For firms associated with paper and printing, we find the weakest effect of distance. Figure 9 shows the corresponding estimates for the effect of *Industry Share %*. With two exceptions (wood products and construction), the point estimates are significant at the 5%-level and positive. The effect is particularly strong for paper and printing, suggesting a strong co-agglomeration pattern in this industry. Overall, the positive effects of *Industry Share %* confirm the view that firms tend to choose locations where there are particularly many other firms of their own industry (Ellison et al., 2010), but it does not provide an alternative explanation for the home-advantage effect.

²⁸ See Appendix B6 for additional information on the construction of the data.

[Insert Figure 8 and 9 here]

6.6 Other Alternative Explanations

Apart from external agglomerations, the firms' location choice could be affected by "internal agglomerations", which describe advantages that occur through close distance to existing plants or branches within a company (Alcácer and Delgado, 2016). If a company had already plants or subsidiaries in West Germany, the location choice might have been influenced by the location of these plants. The latter holds also for East German firms that were themselves subsidiaries of West German firms. If the West German operations of an East German firm, whether headquarters or branches, happened to be located near the later inner-German border, it is conceivable that the decision to locate west of the inner-German border was not caused by home advantages but by internal agglomeration advantages. That is why the home-advantage effect could be biased. However, given that most relocated firms were privately-owned (see Table 1), and therefore, on average, rather small and medium sized firms without any subsidiaries, this concern should be of less relevance in our setting. When we run the main regression model separately for privately-owned firms (excluding corporations and limited liability companies), the results hardly change compared to the full-sample results (see Table A4 in Appendix A3).

Chain migration is another alternative explanation for our findings. We use this term for decision makers that did not choose a new location independently, based on a rational comparison of alternatives, but by imitating the choice of the neighbouring firms. Given that we only observe the outcome of all relocation decisions but have no micro-level data on the individual decision processes, we cannot directly quantify the role of chain migration. The mere observation that several East German companies of the same origin chose the same new location in the West should not be used to infer chain migration, because the companies involved may have independently identified the new location as optimal. Take for example the case of the Thuringian city of Sonneberg mentioned above. Several of this city's firms relocated to the near-by city of Coburg in the north of Bavaria which can just as easily be attributed to independently identified home advantages as to chain migration. Even more complicated is the case of Hamburg where many firms from places in the north of East Germany (e.g., from Schwerin in Mecklenburg) moved to. No chain migration is needed to explain this pattern. Based on our model, Hamburg could have been the optimal choice for many East German companies due to both its short distance and its agglomeration advantages. Identifying chain migration in a narrow sense is challenging.

In Appendix A4, we take a closer look at the location decision of East German firms of the same origin. While chain migration may have played a role in individual cases, we find no convincing evidence for systematic chain migration.

7 Conclusion

This paper studies the determinants of firms' location choice by using an exogenous shock: the German division after World War II that led many established East German firms to search for a new location in West Germany. Based on a newly constructed data set, we show that the East German decision-makers, whether firm owners or managers, preferred West German counties that were close to their former home markets compared to counties that were farer away, and they preferred counties that were part of a larger regional agglomerations with highly productive firms to populous counties with less productive firms and with a location at the periphery, all other things equal. The home-advantage effect was particularly strong when the East German firm originated from an area close to the inner-German border, while it was weaker when firms came from areas more in the east, yet still significant. By analyzing the relocation choice of firms of different industries, we find heterogenous results, suggesting that the home-advantage effect was less strong for firms in world-market oriented industries such as electrical engineering compared to industries in which firms were more focused on local markets such as the manufacturers of wooden products and foodstuff.

Previous research has concentrated on FDI decisions of multinational companies, when analyzing the location choice of firms. However, multinational companies represent, on average, only the very upper tail of the firm distribution of a country. By their nature, these firms are typically larger and more export-oriented than the national average. As suggested by our model, the location choice of multinationals should thus be more affected by positive agglomeration externalities when compared with the location choice of the "average" firm, which is not only smaller but also less export oriented. Since this paper relies on a far more representative set of companies of different industries and sizes, also including firms that were more oriented on local markets, our results should provide less biased estimates for the determinants of firms' location choices.

The results of this paper do not contradict but specify the findings of the literature on the positive externalities of cities and agglomerations. In particular, our findings support the often-held assumption that founders seldom leave their own home territory because they can draw on proven networks and use a lot of experience there. Because many founders live in

agglomerations for demographic reasons alone, this “home bias” reinforces the growth and persistence of existing agglomerations. Only massive exogenous shocks may change these patterns. In West Germany, some underdeveloped regions may have benefited from the historical coincidence that the inner-German border was established near them. Many East German firms just “jumped” across the border to keep their home advantages. This relocation of firms lead to an unexpected inflow of entrepreneurship, human capital, and technological knowledge that might have fostered economic development in the receiving Western counties.

References

- Alcácer, Juan and Mercedes Delgado (2016). Spatial Organization of Firms and Location Choices Through the Value Chain. *Management Science* 62(11), 3213-3234.
- Bähr, Johannes (1997). Die Firmenabwanderung aus der SBZ/DDR und aus Berlin-Ost. In: Wolfram Fischer (ed.). *Wirtschaft im Umbruch: Strukturveränderungen und Wirtschaftspolitik im 19. und 20. Jahrhundert*. Scripta Mercaturae Verlag, St. Katharinen, 229-249.
- Bartik, Timothy J. (1985). Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and the Characteristics of States. *Journal of Business and Economic Statistics* 3(1), 14-22.
- Bauer, Thomas K., Sebastian Braun and Michael Kvasnicka (2013). The Economic Integration of Forced Migrants: Evidence from Post-War Germany. *Economic Journal* 123(571), 99-1024.
- Becker, Sascha O., Karolina Ekholm, Robert Jäckle, and Marc-Andreas Muendler (2005). Location Choice and Employment Decisions: A Comparison of German and Swedish Multinationals. *Review of World Economics* 141(4), 693-731.
- Becker, Sascha O., Lukas Mergele and Ludger Woessmann (2020). *The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism*. Institute of Labor Economics Discussion Paper No. 13032.
- Bikhchandai, Sushil, David Hirshleifer and Ivo Welch (1998). Learning from the Behavior of Others: Conformity, Fads and Informational Cascades. *Journal of Economic Perspectives* 12, 151-170.
- Braun, Sebastian Till and Richard Franke (2021). A County-Level Database on Expellees in West Germany, 1939-1961. *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte* 108(4), 522-540.
- Braun, Sebastian Till, Anica Kramer, Michael Kvasnicka, and Philipp Meier (2021). Local Labor Markets and the Persistence of Population Shocks: Evidence from West Germany. *Journal of Economic Geography* 21(2), 231-260.
- Braun, Sebastian Till and Toman Ohmad Mahmoud (2014). The Employment Effects of Immigration: Evidence from the Mass Arrival of German Expellees in Postwar Germany. *Journal of Economic History* 74(1), 69-108.

- Buenstorf, Guido and Christina Guenther (2010). No Place like Home? Relocation, Capabilities, and Firm Survival in the German Machine Tool Industry after World War II. *Industrial and Corporate Change* 20(1), 1-28.
- Capello, Roberta (2019). Space, Growth and Development: A Historical Perspective and Recent Advances. In: Roberta Capello and Peter Nijkamp (eds). *Handbook of Regional Growth and Development Theories*. 2. Ed., Edward Elgar, Cheltenham, 24-47.
- Carlton, Dennis 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.
- Chung, Wilburg and Juan Alcácer (2002). Knowledge Seeking and Location Choice of Foreign Direct Investment in the United States. *Management Science* 48(12), 1534-1554.
- Ciccone, Antonio and Jan Nimczik (2022). The Long-run Effects of Immigration: Evidence Across a Barrier to Refugee Settlement, Working Paper.
- Davis, Donald R., and David E. Weinstein (2002). Bones, Bombs, and Break Points: The Geography of Economic Activity. *American Economic Review* 92(5), 1269-1289.
- Devereux, Michael P., and Rachel Griffith (1998). Taxes and the Location of Production: Evidence from a Panel of US Multinationals. *Journal of Public Economics* 68(3), 335-367.
- Dittrich, Erich (1951). Der Aufbau der Flüchtlingsindustrien in der Bundesrepublik. *Weltwirtschaftliches Archiv* 67, 327-360.
- Donaldson, Dave and Richard Hornbeck (2016). Railroads and American Economic Growth: A “Market Access” Approach. *Quarterly Journal of Economics* 131(2), 799-858.
- Ellison, Glenn and Edward L. Glaeser (1999): The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?, in: *American Economic Review* 89(2), 311-316.
- Ellison, Glenn, Edward L. Glaeser, and William R. Kerr (2010). What Causes Industry Agglomeration? Evidence from Co-agglomeration Patterns. *American Economic Review* 100(3), 1195-1213.
- Falck, Oliver, Christina Guenther, Stephan Helblich, and William R. Kerr (2013). From Russia with Love: The Impact of Relocated Firms on Incumbent Survival. *Journal of Economic Geography* 13(3), 419-449.
- Figueiredo, Octávio, Paulo Guimarães and Douglas Woodward (2002). Home-field Advantage: Location Decisions of Portuguese Entrepreneurs. *Journal of Urban Economics* 52, 341-361.
- Glaeser, Edward (1998). Are Cities Dying? *Journal of Economic Perspectives* 12, 139-160.
- Glaeser, Edward (2011). *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. Penguin, London.
- Guimarães, Paulo, Octávio Figueiredo and Douglas Woodward (2000). Agglomeration and the Location of Foreign Direct Investment in Portugal. *Journal of Urban Economics* 47, 115-135.
- Hansen, Eric R. (1987). Industrial Location Choice in São Paulo, Brazil: A Nested Logit Model. *Regional Science and Urban Economics* 17, 89-108.
- Head, Keith, and Thierry Mayer (2004). Market Potential and the Location of Japanese Investment in the European Union. *Review of Economics and Statistics* 86(4), 959-972.

- Hecht, Veronika (2017). Location Choice of German Multinationals in the Czech Republic. *Economics of Transition* 25(4), 593-623.
- Hefe, Peter (1998). *Die Verlagerung von Industrie- und Dienstleistungsunternehmen aus der SBZ/DDR nach Westdeutschland*. Franz Steiner Verlag, Stuttgart.
- Held, Colbert C. (1956). Refugee Industries in West Germany after 1945. *Economic Geography* 32(4), 316-335.
- Helpmann, Elhanan, Marcel J. Melitz and Stephen R. Yeaple (2004). Export versus FDI with Heterogeneous Firms. *American Economic Review* 94(1), 300-316.
- Henke, Klaus-Dietmar (2009). *Die amerikanische Besetzung Deutschlands*. 3rd ed. Oldenbourg Verlag, Munich.
- Holtfrerich, Carl-Ludwig (1995). Die Deutsche Bank vom Zweiten Weltkrieg über die Besatzungsherrschaft zur Rekonstruktion 1945-1957. In: Lothar Gall (ed.). *Die Deutsche Bank 1870-1995*. Beck Verlag, Munich, 409-578, 922-952.
- Hornung, Erik (2015). Railroads and Growth in Prussia. *Journal of the European Economic Association* 13(4), 699-736.
- Interessengemeinschaft der in der Ostzone Enteigneten Betriebe (1957). *Westadressen-Verzeichnis der Industrie- und Handelsbetriebe, Banken und Versicherungen aus der sowjetisch besetzten Zone und dem Ostsektor Berlins*. Verlag Klaus Edgar Herfurth, Frankfurt/Main.
- Jacobs, Jane (1970). *The Economy of Cities*. Penguin, London.
- Karlsch, Rainer (1991). Zwischen Partnerschaft und Konkurrenz. Das Spannungsfeld in den Beziehungen zwischen den VEB Filmfabrik Wolfen und der Agfa AG Leverkusen. *Zeitschrift für Unternehmensgeschichte* 36(4), 245-281.
- Krugman, Paul (1991). Increasing Returns and Economic Geography. *Journal of Political Economy* 99(3), 483-499.
- Malmberg, Anders and Peter Maskell (2002). The Elusive Concept of Localization Economies: Towards a Knowledge-Based Theory of Spatial Clustering. *Environment and Planning A* 34(3), 429-449.
- Marshall, Alfred (1890). *Principle of Economics*. Vol. 1, MacMillan, London.
- Maschinenfabrik Buckau R. Wolf AG (eds) (1963). *Ein weites Feld. 125 Jahre Maschinenfabrik Buckau R. Wolf AG*. Hoppenstedts Wirtschafts-Archiv, Darmstadt
- McFadden, Daniel and Kenneth Train (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics* 15(5), 447-470.
- Mirsching, Gerhard (1988). *Vier Jahrzehnte Ingolstädter Automobilbau. Der Weg von DKW und Audi nach 1945*. Bleicher Verlag, Gerlingen.
- Nuppenau, Heino (1974). *Wirkungen der Zonenrandförderung. Analyse der regionalen Einkommenseffekte*. Verlag Weltarchiv GmbH, Hamburg.
- Pascal, Anthony and John McCall (1980). Agglomeration Economies, Search Costs, and Industrial Location. *Journal of Urban Economics* 8, 383-388.
- Peters, Michael (2022). Market Size and Spatial Growth—Evidence from Germany's Post-War Population Expulsions. *Econometrica* 90(5), 2357-2396.

- Redding, Stephen J. and Daniel M. Sturm (2008). The Cost of Remoteness: Evidence from German Division and Reunification. *American Economic Review* 98(5), 1766-1797.
- Redding, Stephen J., Daniel M. Sturm, and Nikolaus Wolf (2011). History and Industry Location: Evidence from German Airports. *Review of Economics and Statistics* 93, 814-831.
- Ritschl, Albrecht and Tamás Vonyó (2014). The Roots of Economic Failure: What Explains East Germany's Falling Behind between 1945 and 1950? *European Review of Economic History* 18, 166-184.
- Salop, Steven C. (1979). Monopolistic Competition with Outside Goods. *Bell Journal of Economics* 10(1), 141-156.
- Schuman, Abel (2014). Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II. *American Economic Journal: Applied Economics* 6(3), 189-205.
- Semrad, Alexandra (2015). *Immigration and Educational Spillovers: Evidence from Sudeten German Expellees in Post-war Bavaria*. Munich Discussion Papers 2015-7.
- Spoerer, Mark and Jochen Streb (2013). *Neue deutsche Wirtschaftsgeschichte des 20. Jahrhunderts*. Oldenbourg Verlag, Munich.
- Steiner, André (2007). *Von Plan zu Plan. Eine Wirtschaftsgeschichte der DDR*. Aufbau Verlagsgruppe, Berlin.
- Wegge, Simone A. (1998). Chain Migration and Information Networks: Evidence from Nineteenth-Century Hesse-Cassel. *Journal of Economic History* 58(4), 957-986.
- Wyrwich, Michael (2020). Migration Restrictions and Long-term Regional Development: Evidence from Large-scale Expulsions of Germans after World War II. *Journal of Economic Geography* 20(2), 481-507.

Tables and Figures

Table 1: Overview of the Number of Relocated Firms

	All firms		In business in 1957	
	#	in %	#	in %
Firms relocated to West Germany	4,676	90.2	3,765	89.3
- AG	343	6.6	263	6.2
- GmbH	395	7.6	312	7.4
- Privately-owned company	3,938	76.0	3,190	75.7
Firms relocated to West Berlin	506	9.8	449	10.6
- AG	75	1.5	73	1.7
- GmbH	71	1.4	65	1.5
- Privately-owned company	360	6.9	311	7.4
Total	5,182	100	4,214	100.0

Note: See the text and Appendix B1 for information on the data source.

Table 2: Descriptive Statistics – Number of Relocated Firms by West German County

	Mean	Std. Dev.	Min.	Max.	Obs.
All Firms	9.12	29.61	0	451	467
AG	0.74	4.01	0	74	467
GmbH	0.82	3.65	0	65	467
Privately-owned	7.57	22.59	0	312	467
Manufacturing	6.37	18.39	0	295	467
Non-Manufacturing	2.76	11.74	0	163	467

Note: This table shows descriptive statistics for the number of firms by West German county (467 counties), including only firms that were in business in 1957. See the text and the Appendix B2 for additional information on the data.

Table 3: Descriptive Statistics – Explanatory Variables

Variable	N	Mean	Std. Dev.	Min	Max
<i>ln(Distance)</i>	1,967,938	5.76467	.419953	.9491327	6.758846
<i>ln(Pop. Den. 1933)</i>	1,967,938	4.872913	1.069821	3.365357	8.476622
<i>ln(Revenues pc 1935)</i>	1,967,938	.2176343	.5374351	-1.278411	2.153847
<i>ln(DMP 1933)</i>	1,967,938	7.429328	.9926441	4.065345	11.07574
<i>Berlin</i>	1,967,938	.0021413	.0462249	0	1
<i>Occupation UK</i>	1,967,938	.3683084	.4823458	0	1
<i>Occupation FR</i>	1,967,938	.1648822	.3710743	0	1
<i>Uni. 1950</i>	1,967,938	.0663812	.2489472	0	1
<i>ln(Distance Uni. 1950)</i>	1,967,938	3.679205	.8833915	0	5.517487
<i>Sea Port</i>	1,967,938	.006424	.0798919	0	1
<i>Highway</i>	1,967,938	.2526767	.4345472	0	1
<i>Latitude</i>	1,967,938	50.39008	1.726426	47.48456	54.80481
<i>Longitude</i>	1,967,938	9.40001	1.769666	6.107838	13.71066
<i>Border East</i>	1,967,938	.0877944	.2829958	0	1
<i>Tax Rate</i>	1,192,562	2.900547	.3025944	2.1	4.375
<i>Expellees 1950</i>	1,963,724	18.85881	9.293584	2.640405	44.1088

Note: This table shows descriptive statistics for all explanatory variables used in the main model. The variables are defined as follows: *ln(Distance)* is the natural logarithm of the great-circle distance (in km) between the old location of firm *i* and county *c*; *ln(Pop. Den. 1933)* is the natural logarithm of the population density of county *c* in 1933; *ln(Revenues pc 1935)* is the natural logarithm of the per-capita firm revenues in county *c* in 1935; *ln(DMP 1933)* is the natural logarithm of the Domestic Market Potential in 1933, as defined in the text; *Berlin* is a dummy variable =1 if county *c* is West Berlin and 0 for all other countries; *Occupation UK* [*Occupation FR*] is =1 if county *c* was part of the British [French] occupation zone, and 0 otherwise (with the US occupation zone as reference category); *Uni. 1950* is =1 if a University was located in county *c* in 1950; *ln(Distance Uni. 1950)* is the natural logarithm of (1 + the minimum distance (in km) to the next university); *Sea Port* =1 for three major sea port cities (Bremen, Hamburg, and Emden), and 0 otherwise; *Highway* =1 if a highway crossed county *c* in 1950, and 0 otherwise; *Latitude* [*Longitude*] indicates the latitude [longitude] of county *c* based on its geographic centroid; *Tax Rate* is the average business tax rate factor in county *c*; *Expellees 1950* is the share (in %) of expellees from former East German territories relative to the population in county *c* in 1950. The number of observations is smaller for the variable *Expellees 1950* because data for West Berlin is not available. Data for *Tax Rate* is also not available for all counties. See the Appendix B3-B5 for more information on the data and its sources.

Table 4: Mixed Logit Choice Model – Main Results

	(1) Choice	(2) Choice	(3) Choice	(4) Choice	(5) Choice	(6) Choice	(7) Choice	(8) Choice
<i>ln(Distance)</i>	-1.025*** (0.016)	-0.966*** (0.017)	-1.291*** (0.024)	-1.233*** (0.028)	-1.189*** (0.028)	-1.150*** (0.028)	-1.207*** (0.031)	-1.191*** (0.031)
<i>ln(Pop. Density 1933)</i>	0.729*** (0.010)	0.363*** (0.016)	0.356*** (0.016)	-0.093*** (0.029)	-0.063** (0.028)	-0.062** (0.030)	-0.072** (0.030)	-0.071** (0.030)
<i>ln(Revenue pc 1935)</i>		1.279*** (0.039)	1.177*** (0.038)	1.057*** (0.040)	0.940*** (0.042)	0.832*** (0.044)	0.858*** (0.044)	0.872*** (0.044)
<i>ln(DMP 1933)</i>			0.178*** (0.010)	0.737*** (0.028)	0.627*** (0.029)	0.602*** (0.030)	0.549*** (0.032)	0.539*** (0.032)
<i>Berlin</i>				6.526*** (0.297)	5.355*** (0.320)	5.278*** (0.330)	5.025*** (0.338)	5.115*** (0.342)
<i>British Occupation</i>				-0.213*** (0.036)	-0.082** (0.037)	-0.172*** (0.040)	-0.160** (0.069)	-0.135** (0.069)
<i>French Occupation</i>				-0.081 (0.064)	-0.114* (0.064)	-0.065 (0.066)	-0.205*** (0.072)	-0.227*** (0.073)
<i>University 1950</i>					0.613*** (0.080)	0.373*** (0.083)	0.409*** (0.083)	0.393*** (0.083)
<i>ln(Distance to University 1950)</i>					0.000 (0.029)	-0.048* (0.029)	-0.059** (0.029)	-0.078*** (0.029)
<i>Sea Port</i>						0.903*** (0.065)	1.119*** (0.084)	1.226*** (0.087)
<i>Highway</i>						0.214*** (0.040)	0.219*** (0.040)	0.206*** (0.040)
<i>Latitude</i>							-0.050** (0.023)	-0.077*** (0.024)
<i>Longitude</i>							-0.072*** (0.017)	-0.094*** (0.018)
<i>Eastern Border</i>								0.352*** (0.064)
N	1967938	1967938	1967938	1967938	1967938	1967938	1967938	1967938
Counties	467	467	467	467	467	467	467	467
Firms	4214	4214	4214	4214	4214	4214	4214	4214
Sample	All	All	All	All	All	All	All	All

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. Standard errors are in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and Table 3 for variable descriptions.

Table 5: Mixed Logit Choice Model – Main Results with Further Controls

	(1) Choice	(2) Choice	(3) Choice	(4) Choice
<i>ln(Distance)</i>	-1.218*** (0.035)	-1.218*** (0.035)	-1.440*** (0.043)	-1.468*** (0.043)
<i>ln(Pop. Density 1933)</i>	-0.044 (0.031)	-0.048 (0.031)	-0.094*** (0.031)	-0.013 (0.034)
<i>ln(Revenue pc 1935)</i>	0.736*** (0.050)	0.737*** (0.050)	0.871*** (0.045)	0.805*** (0.046)
<i>ln(DMP 1933)</i>	0.554*** (0.034)	0.554*** (0.034)	0.564*** (0.033)	0.629*** (0.035)
<i>Tax Rate</i>		-0.162** (0.074)		
<i>Expellees 1950</i>				0.030*** (0.005)
N	1075117	1075117	1754490	1754490
Counties	283	283	466	466
Firms	3799	3799	3765	3765
Sample	Tax Data Available	Tax Data Available	Excl. W-Berlin	Excl. W-Berlin
Controls	Yes	Yes	Yes	Yes

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. In columns (1) and (2), we include only counties for which business tax data is available. In columns (3) and (4), we exclude West Berlin. We include the full set of controls that we use in column (8) of Table 4 in all columns, but we only display the estimated coefficients of *ln(Distance)*, the various agglomeration proxies, and the variables *Tax Rate* and *Expellees 1950*. Standard errors are in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and Table 3 for variable descriptions.

Table 6: Mixed Logit Choice Model – Interpretation of Coefficients

Variable	Coefficient	OR	Std. Dev.	Change in %
<i>ln(Distance)</i>	-1.191	0.304	0.42	-29.4
<i>ln(Pop. Density 1933)</i>	-0.071	0.932	1.07	-7.3
<i>ln(Revenues pc 1935)</i>	0.872	2.391	0.54	75.1
<i>ln(DMP 1933)</i>	0.539	1.714	1.05	74.6

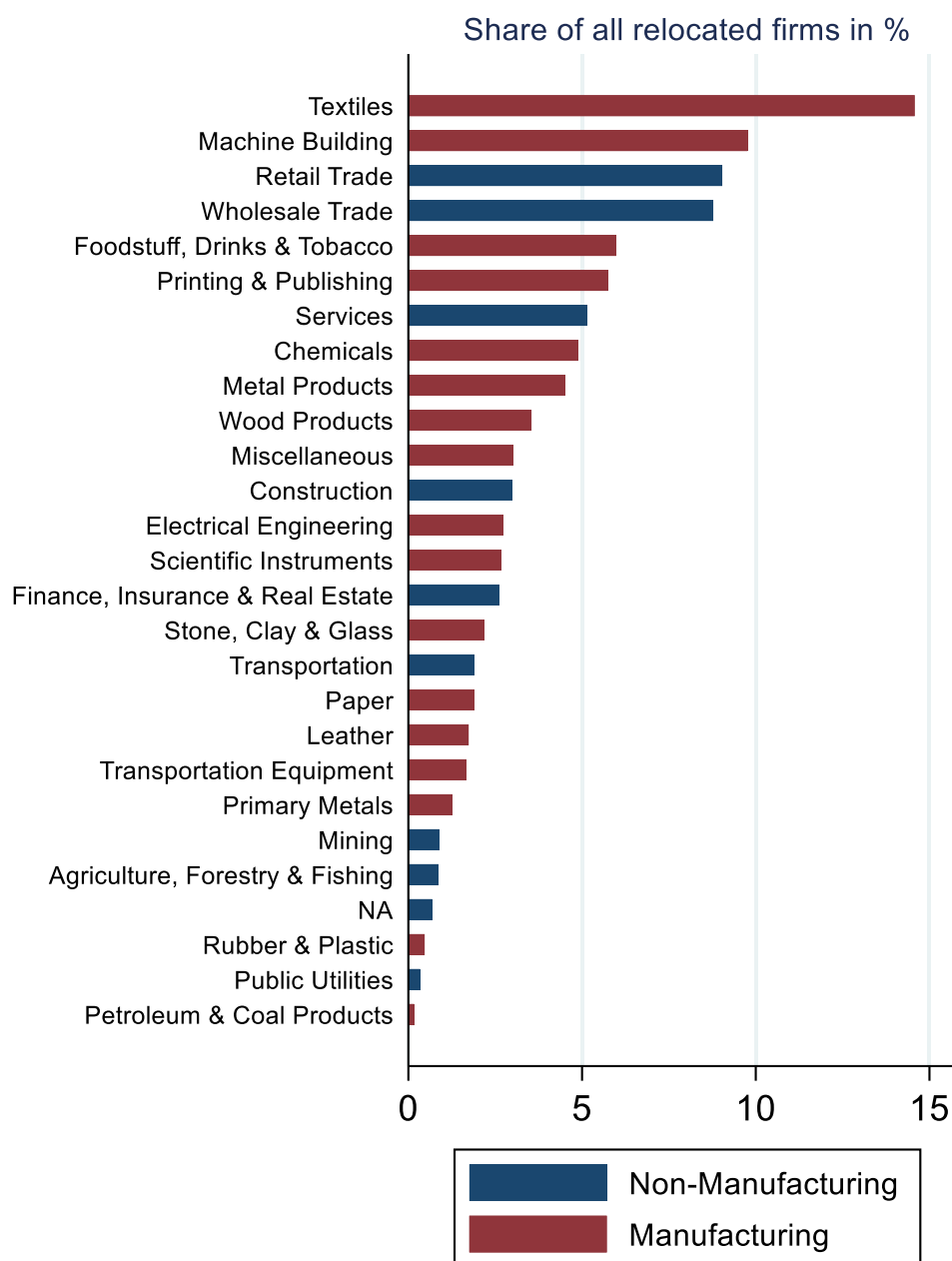
Note: This tabel provides an interpretation of the economic magnitude of the estimated coefficients. The coefficients are from the main model in column (8) of Table 4. We transform these coefficients into odds ratios (OR). Change in % is the change in the odds of choosing a county when we increase the respective variable by one standard deviation. All figures are rounded. See the text and Table 3 for variable descriptions.

Table 7: Mixed Logit Choice Model – Sample Split Tests

	(1) Choice	(2) Choice	(3) Choice	(4) Choice	(5) Choice	(6) Choice	(7) Choice
ln(Distance)	-1.191*** (0.031)	-2.360*** (0.111)	-1.043*** (0.033)	-1.675*** (0.062)	-1.069*** (0.044)	-0.947*** (0.045)	-1.590*** (0.050)
ln(Pop. Density 1933)	-0.071** (0.030)	-0.185* (0.098)	-0.065** (0.032)	-0.069* (0.040)	-0.095** (0.047)	0.024 (0.050)	-0.153*** (0.038)
ln(Revenue pc 1935)	0.872*** (0.044)	0.598*** (0.140)	0.903*** (0.047)	0.843*** (0.058)	0.919*** (0.070)	1.037*** (0.075)	0.788*** (0.056)
ln(DMP 1933)	0.539*** (0.032)	0.551*** (0.104)	0.546*** (0.034)	0.555*** (0.042)	0.536*** (0.050)	0.552*** (0.053)	0.542*** (0.041)
N	1967938	197541	1770397	1069430	898508	836397	1131541
Counties	467	467	467	467	467	467	467
Firms	4214	423	3791	2290	1924	1791	2423
Sample	All	Old Places at the Border	Old Places NOT at the Border	Western Troops in 1945	NO Western Troops in 1945	Large Cities	Excl. Large Cities
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. In columns (2) and (3), we split the sample such that column (2) only includes firms from old places in grid cells close to the border and column (3) all firms from places not in border grid cells. In columns (4) and (5), we split the sample such that column (3) only includes firms from old places in territories that were temporarily occupied by the Western Allies in 1945 and column (4) includes firms from places in territories that were not temporarily occupied by the Western Allies in 1945. In columns (6) and (7), we split the sample such that column (6) only includes firms from cities with a pre-war size of more than 100,000 inhabitants, and column (7) includes all other firms. In all columns we include the full set of controls that we use in column (8) of Table 4, but we only display the estimated coefficients of *ln(Distance)*, *ln(Pop. Density 1933)*, *ln(Revenue pc 1935)*, and *ln(DMP 1933)*. Standard errors are in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and the Appendix B2 for more information on the data and sample definition.

Figure 1: Relocated Firms by Industries



Notes: This figure shows the share of firms by industries (in %) relative to all relocated firms that were in business in 1957. "NA" describes firms without information on the related industry. See Appendix B1 for more information on the industry definitions.

Figure 2: Distribution of Relocated Firms in West Germany

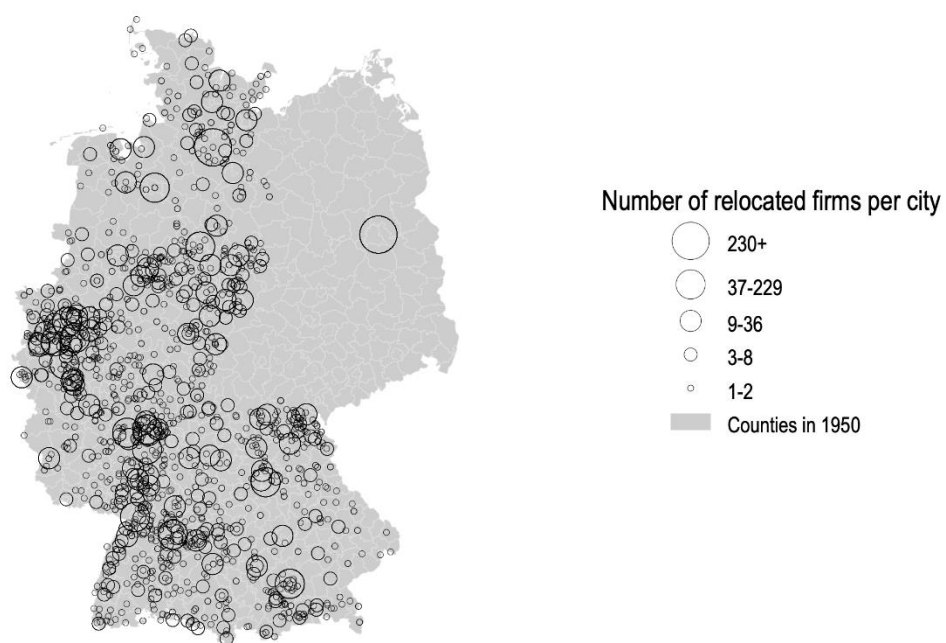


Figure 3: Distribution of Firms' Old Locations in East Germany

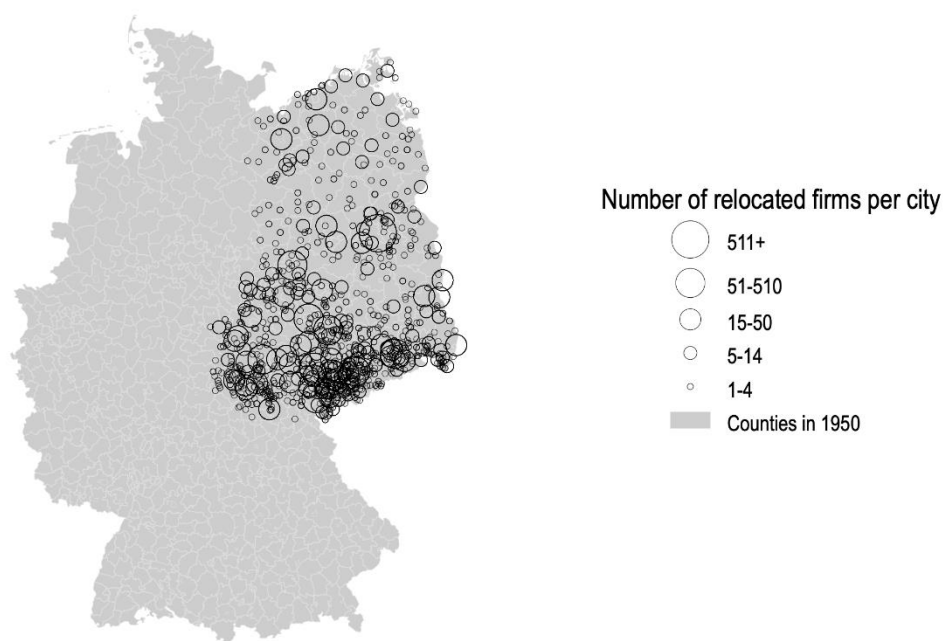
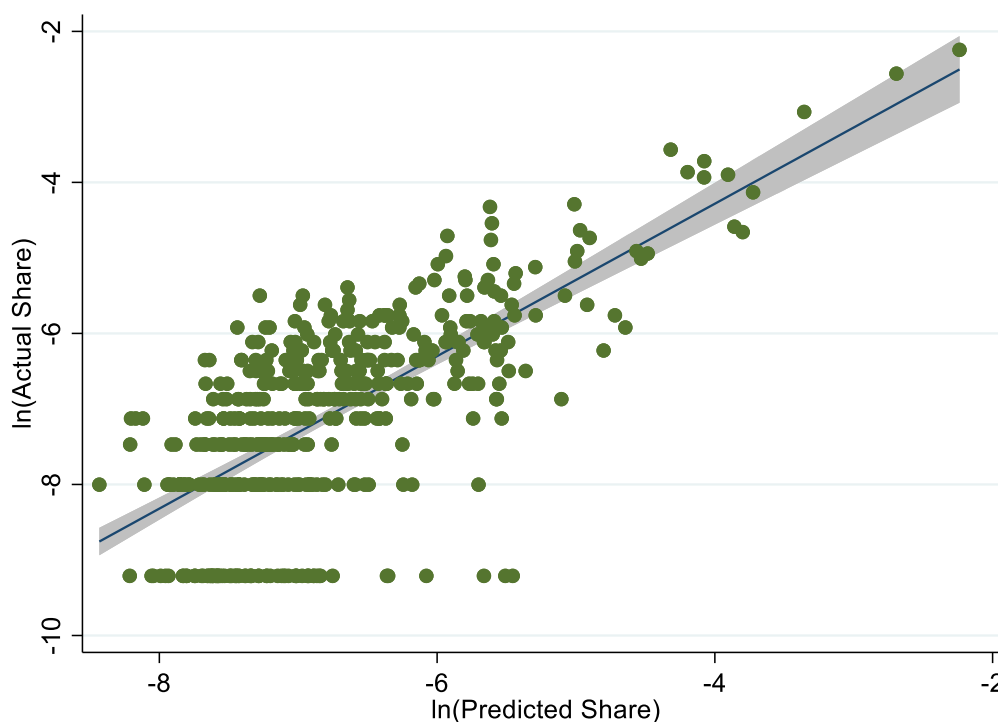
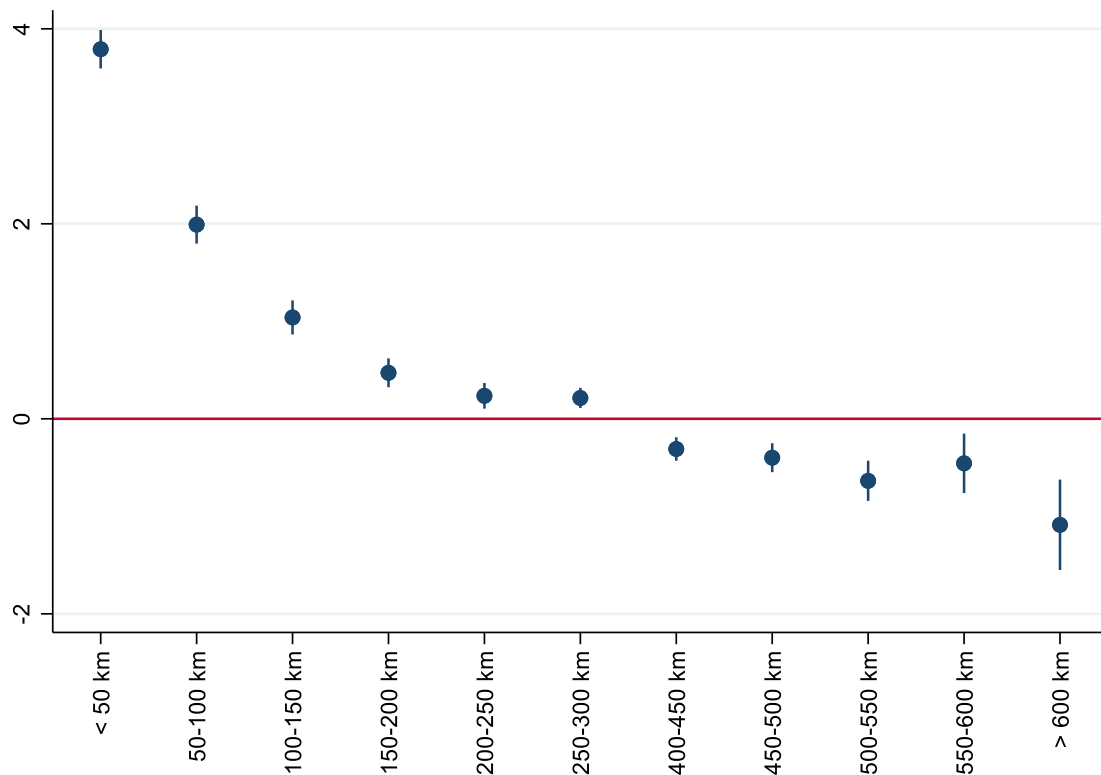


Figure 4: Predicted Share of Relocated Firms by County



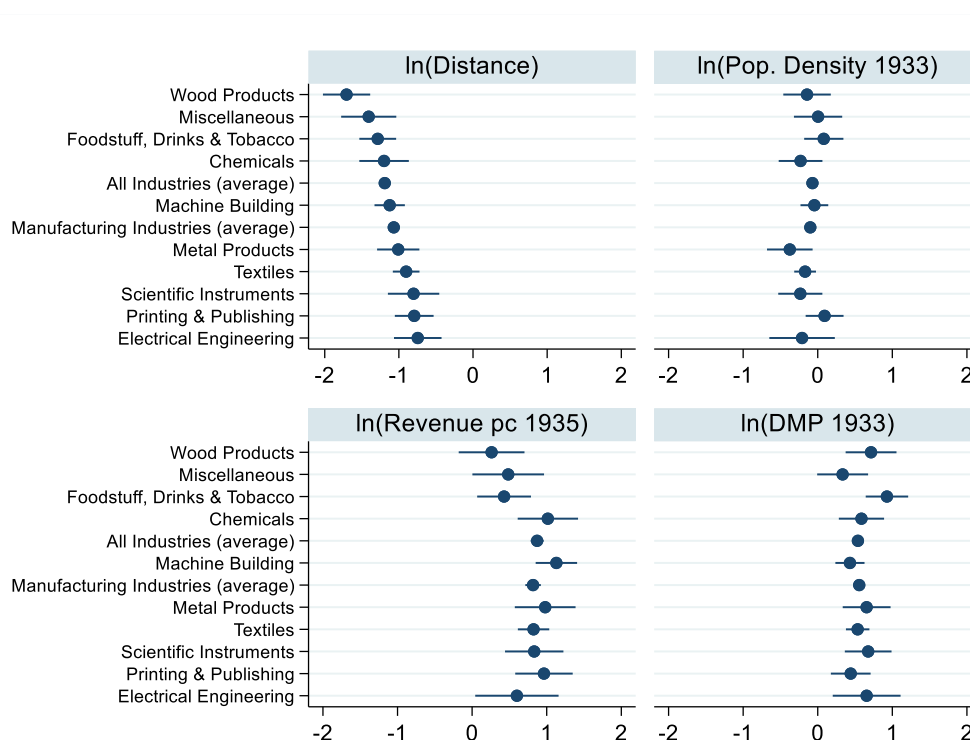
Notes: This figure compares the predicted shares with the actual (historical) shares of firms relocated to a county relative to the total number of relocated firms. We compute the predicted shares based on the full-sample model with all controls (column (8) of Table 4). The natural logarithm of the predicted shares is plotted on the x-axis and natural logarithm of the actual shares on the y-axis. In both cases, we add 0.0001 to the respective shares before taking the logarithm, since there are counties for which the actual shares are zero. The line indicates the fitted values and the grey area the 95% confidence interval. When running the univariate OLS regression $\ln(\text{Actual Share}) = \beta \ln(\text{Predicted Share})$, the estimated $\hat{\beta}$ is 1.0449 (standard error: 0.0064) and the R^2 is 0.98.

Figure 5: Mixed Logit Choice Model – Non-Linear Distance Effects



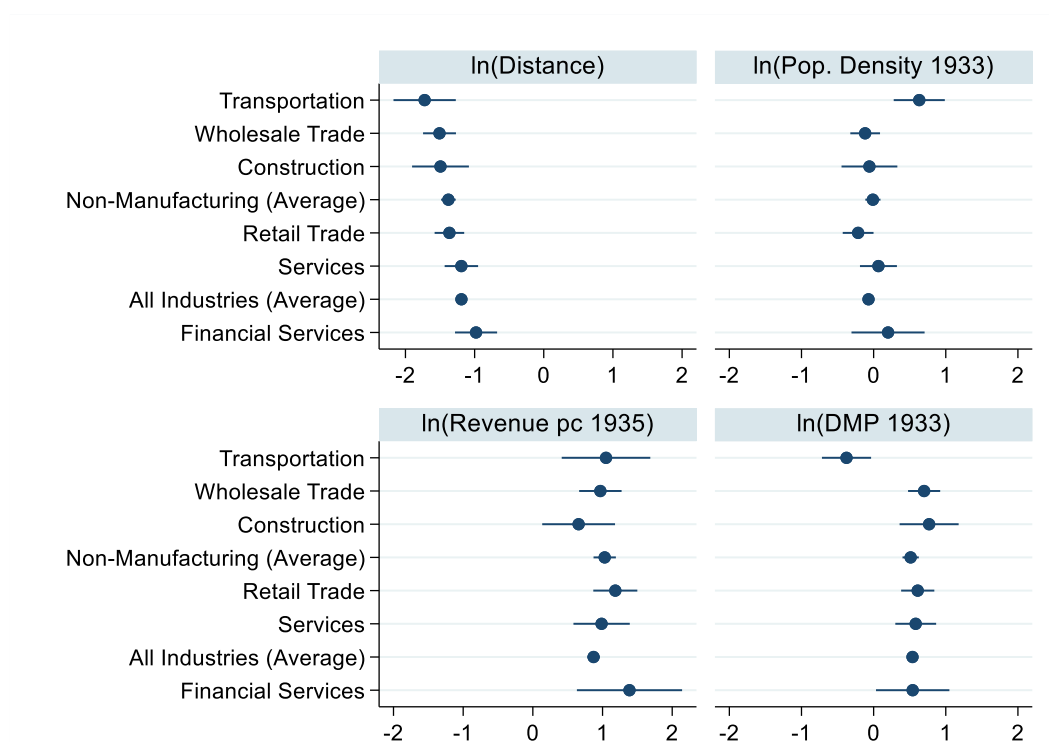
Notes: This figure shows the estimated coefficients for different distance dummies using a mixed logit choice model. The endogenous variable is *Choice*, which is =1 if a firm selected the respective county, and 0 otherwise. We use the full set of agglomeration proxies and controls as in column (8) of Table 4. The dots show the point estimates (with standard errors) for the coefficients of each dummy variable, indicating whether the distance between the old firm location and the potential county of choice is in the respective interval. The reference category is the median-distance interval (300-400 km). See the text and the appendix for additional information on the data.

Figure 6: Mixed Logit Choice Model – Differences across Manufacturing Industries



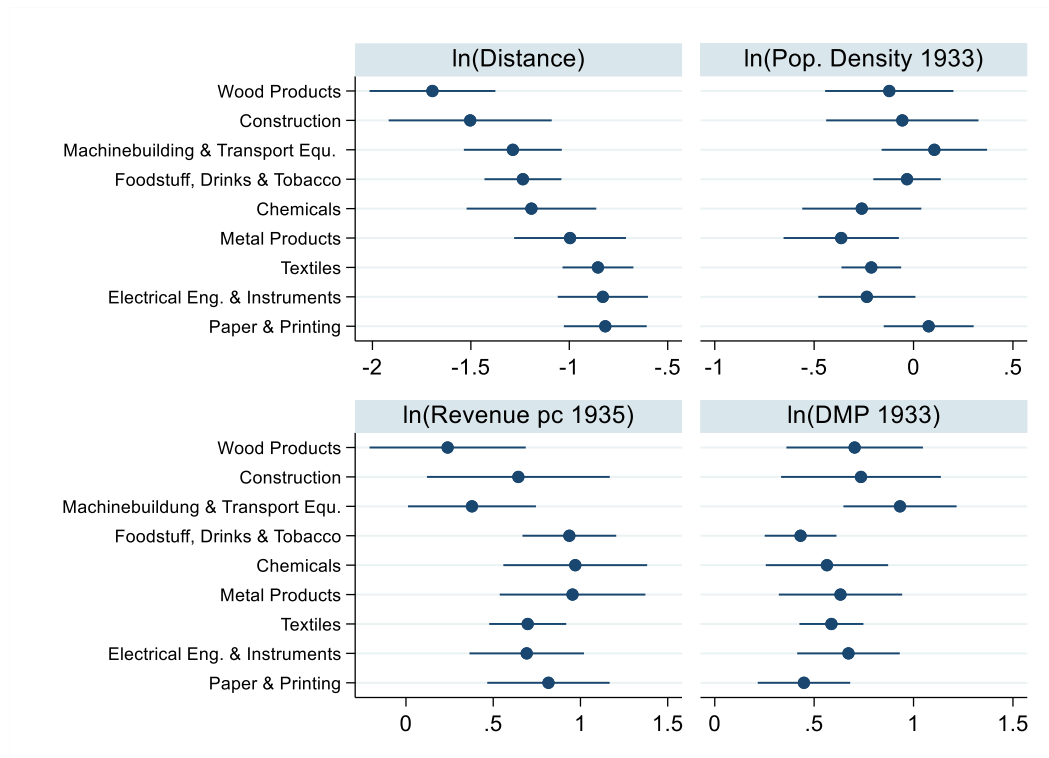
Notes: This figure shows the estimated coefficients for the effect of $\ln(\text{Distance})$, $\ln(\text{Pop. Density 1933})$, $\ln(\text{Revenue pc 1935})$, and $\ln(\text{DMP 1933})$ by manufacturing industries, respectively. The endogenous variable is *Choice*, which is =1 if a firm selected the respective county, and 0 otherwise. We run the baseline model with all controls (see column (8) of Table 4) but include only firms of a specific industry. We report the coefficients for the ten largest manufacturing industries (based on the number of firms in the data set) as well as the average over all industries and the average over all manufacturing industries for comparison. The dots show the point estimates (with standard errors) for the coefficients. We order the industries according to the size of the effect of $\ln(\text{Distance})$ in all panels. See the text and the Appendix for additional information on the data.

Figure 7: Mixed Logit Choice Model – Differences across Non-Manufacturing Industries



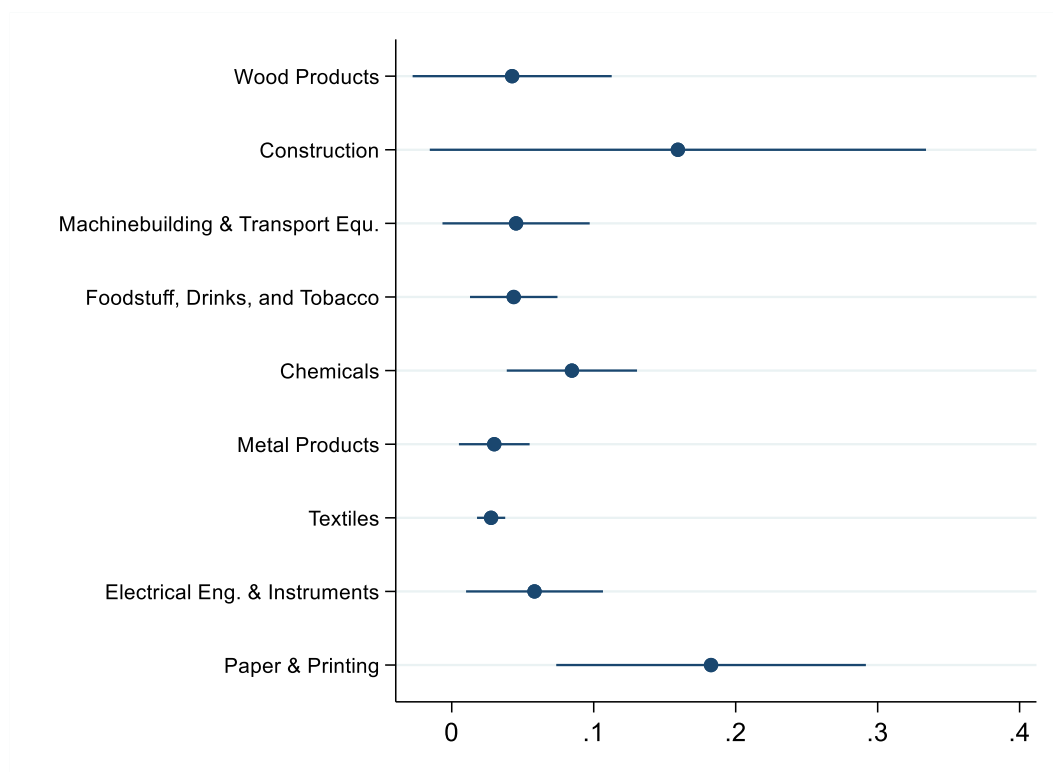
Notes: This figure shows the estimated coefficients for the effect of $\ln(\text{Distance})$, $\ln(\text{Pop. Density 1933})$, $\ln(\text{Revenue pc 1935})$, and $\ln(\text{DMP 1933})$ by non-manufacturing industries, respectively. The endogenous variable is *Choice*, which is =1 if a firm selected the respective county, and 0 otherwise. We run the baseline model with all controls (see column (8) of Table 4) but include only firms of a specific industry. We report the coefficients for the six largest non-manufacturing industries (based on the number of firms in the data set) as well as the average over all industries and the average over all non-manufacturing industries for comparison. The dots show the point estimates (with standard errors) for the coefficients. We order the industries according to the size of the effect of $\ln(\text{Distance})$ in all panels. See the text and the Appendix for additional information on the data.

Figure 8: Mixed Logit Choice Model – Main Results with Industry-Share Controls



Notes: This figure shows the estimated coefficients for the effect of $\ln(\text{Distance})$, $\ln(\text{Pop. Density 1933})$, $\ln(\text{Revenue pc 1935})$, and $\ln(\text{DMP 1933})$ by different industries, respectively. The endogenous variable is *Choice*, which is =1 if a firm selected the respective county, and 0 otherwise. We include only firms of a specific industry and run the baseline model with all controls (see column (8) of Table 4) and *Industry Share %* as additional control, respectively. *Industry Share %* is the pre-war employment share (in %) of the respective industry relative to the total number of employees in the respective county. We report the coefficients for industries with at least 100 firms in the data set and for which employment data is available to compute the variable *Industry Share %*. The dots show the point estimates (with standard errors) for the coefficients. We order the industries according to the size of the effect of $\ln(\text{Distance})$ in all panels. See the text and the Appendix for additional information on the data.

Figure 9: Mixed Logit Choice Model – Effect of Industry Share on Location Choice



Notes: This figure shows the estimated coefficients for the effect of *Industry Structure %* by different industries, respectively. The endogenous variable is *Choice*, which is =1 if a firm selected the respective county, and 0 otherwise. We include only firms of a specific industry and run the baseline model with all controls (see column (8) of Table 4) and *Industry Share %* as additional control, respectively. *Industry Share %* is the pre-war employment share (in %) of the respective industry relative to the total number of employees in the respective county. We report the coefficients for industries with at least 100 firms in the data set and for which employment data is available to compute the variable *Industry Share %*. The dots show the point estimates (with standard errors) for the coefficients. We order the industries according to the size of the effect of $\ln(\text{Distance})$ (see Figure 8). See the text and the appendix for additional information on the data.

Appendix

Appendix A – Additional Statistics

A1 Agglomeration Proxies based on 1950 Census Data

Throughout the paper, we have used data from the 1930s to construct proxies of agglomeration. While we prefer data from 1933 to avoid any endogeneity concerns, we show in this section that the data are highly correlated, suggesting a relatively strong persistence of agglomerations, and that, consequently, the results look similar when using post-war data from the West German census of 1950 to test the determinants of location choice.

Table A1 shows descriptive statistics for all three agglomeration proxies that we constructed with data from the 1950 census. Using the same notation as in the paper, $\ln(\text{Pop. Den. } 1950)$ is the natural logarithm of the population density in a county in 1950, $\ln(\text{Revenues pc } 1950)$ is the natural logarithm of the firm revenues in a county in 1950 divided by the population, and $\ln(\text{DMP } 1950)$ is the natural logarithm of the Domestic Market Potential (as defined in section 5.2). We also computed the differences of the logged population densities in 1933 and 1950, $\Delta \ln(\text{Pop. Density})$, and the logged firm revenues per capita, $\Delta \ln(\text{Revenues pc})$.

Table A1: Descriptive Statistics – Agglomeration Proxies (1950 Census Data)

Variable	N	Mean	Std. Dev.	Min	Max
$\ln(\text{Pop. Den. } 1950)$	1,963,724	5.193806	.9737158	3.733483	8.325982
$\ln(\text{Revenues pc } 1950)$	1,963,724	8.153758	.6404219	6.848005	10.12371
$\ln(\text{DMP } 1950)$	1,963,724	7.753298	.9004341	4.411936	10.99784
$\Delta \ln(\text{Pop. Density})$	1,963,724	.3286261	.2068276	-.2796793	1.693291
$\Delta \ln(\text{Revenues pc})$	1,963,724	7.938973	.4034857	6.236334	9.75561

Note: This table shows descriptive statistics for agglomeration proxies constructed with 1950 census data. The variables are defined as follows: $\ln(\text{Pop. Den. } 1950)$ is the natural logarithm of the population density of county c in 1950; $\ln(\text{Revenues pc } 1950)$ is the natural logarithm of the per-capita firm revenues in county c in 1950; $\ln(\text{DMP } 1950)$ is the natural logarithm of the Domestic Market Potential in 1933, as defined in the text; $\Delta \ln(\text{Pop. Density})$ [$\Delta \ln(\text{Revenues pc})$] is the change of the natural logarithm of the population densities, $= \ln(\text{Pop. Den. } 1950) - \ln(\text{Pop. Den. } 1933)$ [revenues per capita, $= \ln(\text{Revenues pc } 1950) - \ln(\text{Revenues pc } 1933)$]; Note that data for West Berlin is not available. See Appendix B4 for more information on the data and its sources.

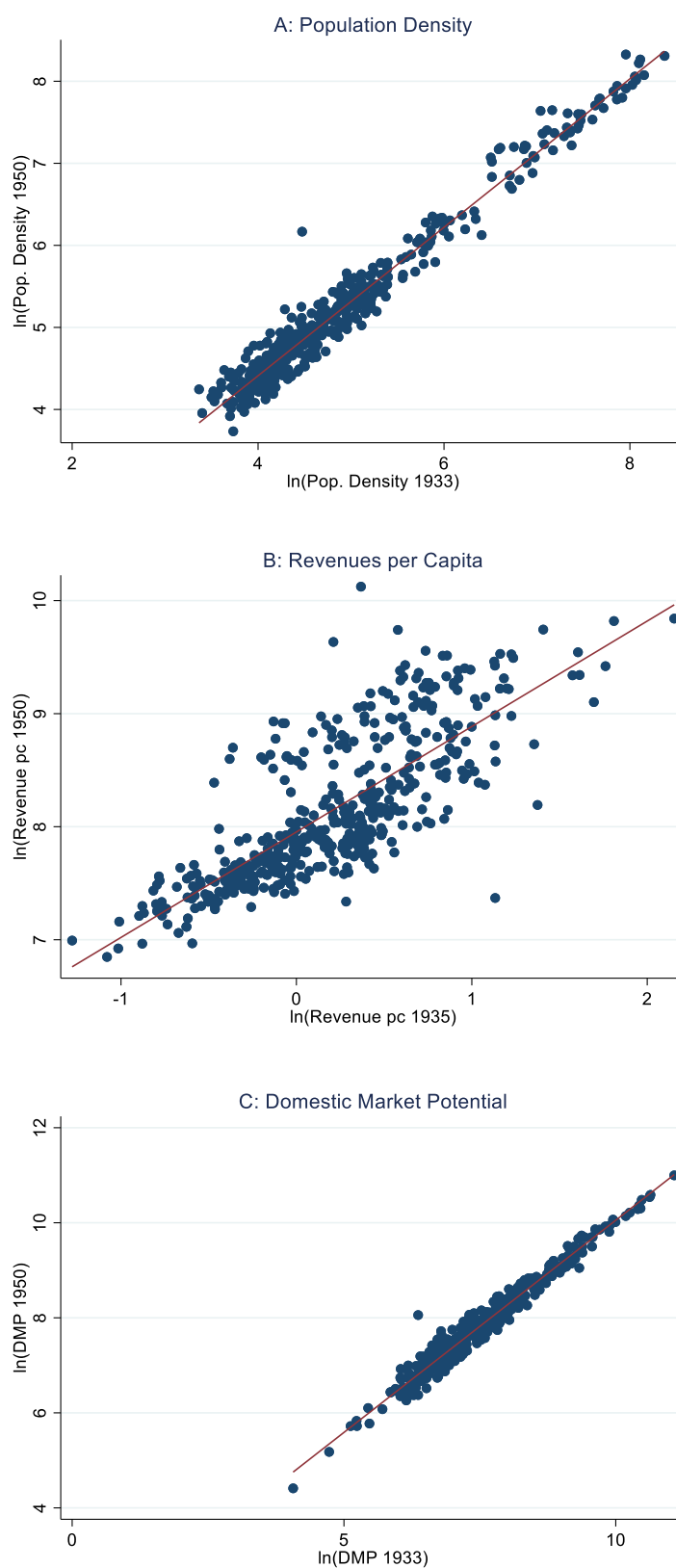
In Figure A1, we compare all three agglomeration proxies when measured with 1933 and 1950 data, respectively. Panel A shows the correlation between the logged county-level population densities. The values for 1950 are plotted on the y-axis, the corresponding values for 1933 on the x-axis. Despite of large population movements that resulted from World War II, including

the forced migration of the Eastern German population, the agglomeration pattern looks quite similar in 1950 when compared to 1933. Panel B shows the correlation of the logged firm revenues per capita, for which we find a stronger dispersion, suggesting a weaker persistence compared to the population density. Last, in Panel C, we show the correlation of the logged Domestic Markt Potential. The latter proxy has the highest correlation.

To test whether our results are affected by the changes in the county population, we now use additional data from the 1950 census. In columns (1) and (2) of Table A2, we show for comparison the estimated coefficients for $\ln(\text{Distance})$ and all three agglomeration proxies based on 1933/35 data in the model with all controls for the full sample and for a sample without West Berlin. In column (3), we estimate the same model as in column (2) but with agglomeration proxies based on 1950 census data. Note that we had to exclude West Berlin since data on firm revenues are not available for the year 1950. The effect of $\ln(\text{Distance})$ remains almost unchanged compared to the specification with 1933/35 data. However, when comparing the coefficients for the effect of population density, we find a reversal in the direction of the effect. While the coefficient of $\ln(\text{Pop. Density } 1933)$ is significantly negative, it is significantly positive for $\ln(\text{Pop. Density } 1950)$. By contrast, we find a weaker effect of revenues per capita in column (3), while the effect of the Domestic Market Potential is almost unchanged. In column (4), we include the agglomeration proxies based on 1933/35 data as well as the change of the natural logarithm of the population density and revenues per capita between 1933/35 and 1950, respectively. The effect of $\ln(\text{Distance})$ again hardly changed. By contrast, the effect of $\ln(\text{Pop. Density } 1933)$ became insignificant while the change in population density, $\Delta \ln(\text{Pop. Density})$, is significantly positive. Likewise, we find a significantly positive effect of the change in revenues per capita, $\Delta \ln(\text{Revenue } pc)$. The positive correlation between the relocation choice and the growth in population and economic activity between 1933 and 1950 does not necessarily imply that the East German firms were attracted to places with a booming economy. The opposite may be true: The observable growth in economic activity could be a consequence of the re-settlement of former Eastern firms and their employees.

We conclude that the negative and significant effect of $\ln(\text{Distance})$, which we use to measure the home advantage, is hardly affected by the point in time at which we collect the agglomeration variables. This finding reflects the fact that, despite of large population movements, there was no radical change in the spatial distribution of agglomerations in West Germany after World War II.

Figure A1: Correlation of Agglomeration Proxies in 1933 and 1950



Note: Panel A shows the correlation of the logged population densities in 1950 and 1933, Panel B the correlation of logged revenues per capita, and Panel C the correlation of the logged Domestic Market Potential. The red lines indicate the fitted values.

Table A2: Mixed Logit Choice Model – Alternative Agglomeration Proxies

	(1) Choice	(2) Choice	(3) Choice	(4) Choice
<i>ln(Distance)</i>	-1.191*** (0.000)	-1.440*** (0.000)	-1.444*** (0.000)	-1.492*** (0.000)
<i>ln(Pop. Density 1933)</i>	-0.071** (0.019)	-0.094*** (0.002)		0.042 (0.194)
<i>ln(Revenue pc 1935)</i>	0.872*** (0.000)	0.871*** (0.000)		0.795*** (0.000)
<i>ln(DMP 1933)</i>	0.539*** (0.000)	0.564*** (0.000)		0.656*** (0.000)
<i>ln(Pop. Density 1950)</i>			0.110*** (0.000)	
<i>ln(Revenue pc 1950)</i>			0.514*** (0.000)	
<i>ln(DMP 1950)</i>			0.597*** (0.000)	
<i>Δln(Pop. Density)</i>				1.436*** (0.000)
<i>Δln(Revenue pc)</i>				0.216*** (0.000)
N	1967938	1754490	1754490	1754490
Counties	467	466	466	466
Firms	4214	3765	3765	3765
Sample	All	Excl. W-Berlin	Excl. W-Berlin	Excl. W-Berlin
Controls	Yes	Yes	Yes	Yes

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. In columns (2) to (4), we exclude West Berlin. In all columns we include the full set of controls that we use in column (8) of Table 4, but we only display the estimated coefficients of *ln(Distance)* and the various agglomeration proxies as well as their differences. *p*-values in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and Table 3 and Table A1 for a description of all variables.

A2 Additional Results for the Effect of Distance in Border Grid Cells

In section 6.3, we analyze whether the home advantage, measured with the variable $\ln(\text{Distance})$ is larger for firms that were originally from places close to the West-East border that was established after the war. To test whether the effect of $\ln(\text{Distance})$ differs conditional on the pre-war firm location, we construct samples by subdividing the area of East Germany into grid cells of 0.5 x 0.5 degrees and assign each geo-coded old firm location to the respective grid cells. In addition to the results in Table 7, where we use a “narrow” grid-cell definition, with only one grid cell next to the border, we show additional results for a “wide” grid-cell definition in Table A3. The wide grid-cell definition includes all firms from the sample of the narrow grid-cell definition as well as all firms that are in all neighbouring grid cells. Using the wide definition, the number of firms included in the border sample increases to 1,743.

In Table A3, we show for comparison the baseline results (all firms) in column (1) and the sample split regressions with the narrow grid-cell definition in columns (2) and (3) (as shown in Table 7). Column (4) reports the new results for the border sample with the wide grid-cell definition and column (5) shows the results for all other firms, which are not included in column (4). When using the wide definition, we still find a significantly negative effect of $\ln(\text{Distance})$, but the effect is smaller than in the narrow definition used in column (2). By contrast, the effect of $\ln(\text{Revenue pc 1935})$ is much larger. These results show that the home advantage, which we measure with $\ln(\text{Distance})$ decreases when we include additional firms from areas that were more distant from the border, supporting the results shown in section 6.2, where we argue that the effect of distance decreased with rising distances. By contrast, for firms that were more far away from the border, suggesting fewer pre-war links to customers or suppliers in the West, agglomerations with high productivity, measured with population density and firm revenues per capita, were relatively more attractive.

Table A3: Mixed Logit Choice Model – Additional Sample Split Test

	(1) Choice	(2) Choice	(3) Choice	(4) Choice	(5) Choice
<i>ln</i> (Distance)	-1.191*** (0.031)	-2.360*** (0.111)	-1.043*** (0.033)	-1.832*** (0.063)	-0.993*** (0.040)
<i>ln</i> (Pop. Density 1933)	-0.071** (0.030)	-0.185* (0.098)	-0.065** (0.032)	-0.208*** (0.046)	-0.005 (0.040)
<i>ln</i> (Revenue pc 1935)	0.872*** (0.044)	0.598*** (0.140)	0.903*** (0.047)	0.861*** (0.067)	0.916*** (0.060)
<i>ln</i> (DMP 1933)	0.539*** (0.032)	0.551*** (0.104)	0.546*** (0.034)	0.561*** (0.049)	0.545*** (0.043)
N	1967938	197541	1770397	813981	1153957
Counties	467	467	467	467	467
Firms	4214	423	3791	1743	2471
Sample	All	Old Places at the Border (narrow)	Old Places NOT at the Border (narrow)	Old Places at the Border (wide)	Old Places NOT at the Border (wide)
Controls	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. In column (1), we show the results for the baseline model (see column (8) of Table 4). In columns (2) and (3), we split the sample such that column (2) only includes firms from old places in grid cells close to the border and column (3) all firms from places not in border grid cells. In columns (4) and (5), we split the sample in the same way but use a wider border area. The border sample includes firms from the border grid cells of column (3) and all firms located in all the neighboring grid cells of 0.5 x 0.5 degrees. In all columns, we include the full set of controls that we use in column (8) of Table 4, but we only display the estimated coefficients of *ln*(Distance), *ln*(Pop. Density 1933), *ln*(Revenue pc 1935), and *ln*(DMP 1933). Standard errors are in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and the appendix for more information on the data and sample definition.

A3 Differences across Firm Types

We now test whether the effects differ for the three firm types: corporations (AG), limited liability companies (GmbH), and privately-owned firms. While we have no information about the firm size, we assume that corporations represent, on average, larger firms than limited liability companies, and the latter were, on average, larger than privately-owned firms. These differences might have mattered for the distance effect because smaller firms might have had stronger ties to local markets than large firms. By splitting the sample between corporations and limited liability companies on the one hand and privately-owned firms on the other, we also account for the fact that a subset of the former companies were subsidiaries of larger business groups from outside East Germany. In such cases, the decision where to relocate the firm might have been affected by the location of the headquarter or other affiliated branches of the mother company. By contrast, there is no similar bias for privately-owned firms.

In Table A4, we show estimates for the effects of $\ln(Distance)$ and all agglomeration proxies for different firm types, respectively. For comparison, we report the estimate for the full sample with all controls in the first column. Column (2) includes only corporations (336 firms), column (3) only limited liability companies (377 firms), and column (4) all privately-owned firms (3,501 firms). When compared to the group of privately-owned firms, we find a weaker effect of $\ln(Distance)$ in columns (2) and (3), but the differences are not very large.

Table A4: Mixed Logit Choice Model – Different Type of Firms

	(1) <i>Choice</i>	(2) <i>Choice</i>	(3) <i>Choice</i>	(4) <i>Choice</i>
<i>ln(Distance)</i>	-1.191*** (0.031)	-1.094*** (0.115)	-0.991*** (0.102)	-1.218*** (0.034)
<i>ln(Pop. Density 1933)</i>	-0.071** (0.030)	-0.031 (0.122)	0.092 (0.100)	-0.095*** (0.033)
<i>ln(Revenue pc 1935)</i>	0.872*** (0.044)	1.210*** (0.169)	0.970*** (0.149)	0.836*** (0.048)
<i>ln(DMP 1933)</i>	0.539*** (0.032)	0.724*** (0.124)	0.390*** (0.105)	0.540*** (0.035)
N	1967938	156912	176059	1634967
Counties	467	467	467	467
Firms	4214	336	377	3501
Sample	All	Corporations (AG)	Limited Liability Comp. (GmbH)	Private Companies
Controls	Yes	Yes	Yes	Yes

Notes: This table shows the estimated coefficients for the mixed logit choice model. The endogenous variable is *Choice* in all columns, which is =1 if a firm selected the respective county, and 0 otherwise. In all columns, we include the full set of controls that we use in column (8) of Table 4, but we only display the estimated coefficients of *ln(Distance)* and all three agglomeration proxies. In column (1), we include all firms, column (2) includes only corporations (Aktiengesellschaften), column (3) only limited liability companies (GmbH), and columns (4) all private companies. Standard errors are in parentheses. ***, **, and * denotes significance at the 1%, 5%, and 10% level. See the text and Table 3 for a description of all variables.

A4 Chain Migration

To find out whether chain migration may have played a role in the location decision, we examine here how different the choices were that firms of the same origin made. For example, the 23 firms that left the city of Schwerin in Mecklenburg chose 11 different locations in the West, with a clear focus on Hamburg, which was preferred by 10 firms. Overall, we observe different 2,777 pairs between “old” places in the East and “new” counties in the West. We suspect that decision makers from small towns were more susceptible to chain migration because social ties were stronger, and transparency was greater there than in large cities. To test this hypothesis, we first consider places from which few East German firms came and which we therefore define as “small.”

Table A5 ranks the East German places of origin according to the number of firms that moved from them. In the first line, we report the number of old places for which we observe only one firm that migrated to the West. By assumption, the relocation decisions made in these old places cannot reflect chain migration because there was no second firm to imitate. The second line shows the number of old places for which we observe exactly two migrating firms. Combinations when both of these firms chose the same new place, we interpret as weak evidence for chain migration, because the likelihood of doing so when there are 467 potential options to choose from is rather small. Yet, out of 118 old places with two migrating firms, there were only 37 for which this was the case. This means that the share of combinations with at least two firms is 18.6 percent. We think that this number is no clear proof for chain migration but rather suggests that chain migration was not likely for the remaining 81.4 percent. In the next line, we consider old places with three migrating firms. Here, the share of combinations with at least two firms at a new place is 11.9 percent. The following lines provide similar information for the groups of old places with four and five re-located firms. Given the lower number of places with more than five relocated firms, we continue to report the data for the groups with 6-10 relocated firms, 11-20, 21-30, 31-100, 101-500, and for the group with more than 500 firms. The last group includes only East-Berlin and the group 101-500 the cities of Leipzig, Dresden, Magdeburg, and Chemnitz. For the later groups, the share of combinations with at least two firms in the same West German county is larger than in the earlier groups which already results, among other things, from the higher number of companies relocating. However, most combinations are still combinations for which we observe only one firm in the new place, suggesting again no specific chain migration pattern. This is all the more true because the East German firms companies could have chosen the new locations with at least

two relocated firms independently of each other, for example because of agglomeration advantages or home advantages.

Table A5: Distribution of Relocated Firms by Old Place

Number of Relocated Firms by Old Place	Number of Old Places	Number of Firms	Cum. Number of Firms	Number of combinations with at least two firms	Number of all combinations	Share of combinations with at least two firms in %
1	377	377	377	0.0	377	0.0
2	118	236	613	37	199	18.6
3	67	201	814	21	177	11.9
4	38	152	966	20	127	15.7
5	18	90	1,056	8	80	10.0
6-10	53	387	1,443	47	333	14.1
11-20	40	590	2,033	87	481	18.1
21-30	12	286	2,319	46	202	22.8
31-100	6	299	2,618	50	299	26.7
101 - 500	4	1,008	3,626	153	368	41.6
> 500	1	588	4,214	48	134	35.8

Note: This table provides information about the distribution of relocated firms by old places. We distinguish different groups of old places based on the total number of firms from these places. The first column indicates the groups, starting with places with only one re-located firm in the first line. The second last group (101-500) includes the cities Leipzig, Dresden, Magdeburg, and Chemnitz, the last group (> 500) East-Berlin. The following columns report the number of old places, the number of firms, the cumulative number of firms, the combinations (between an old place in the East and a county in the West) with at least two firm observations, the total number of combinations, and the share of combinations with at least two firms in %.

Appendix B – Data and Variable Description

B1 Relocated Companies

To identify relocated firms, we use the “Register of West addresses of industrial and trade companies, banks, and insurance companies from the Sowjet-occupied zone and the East sector of Berlin”, which was published in 1957 (*Westadressen-Verzeichnis der Industrie- und Handelsbetriebe, Banken und Versicherungen aus der sowjetisch besetzten Zone und dem Ostsektor Berlins*, Verlag Klaus Edgar Herfurth, Frankfurt am Main). This register includes information about the name of the firm, the place where the firm was based in the Soviet zone, the new address in West Germany (and the new firm’s name in case that it changed), and information about the industry affiliation. To illustrate the structure of the original data source, we show the first ten entries, as reported in the register, in Table B1.

Table B1: Register of West Addresses – Structure of the Data

Firm name and former place	Industry	New firm address
<i>Abel</i> Chemnitz	Textilhandel	Abel-Textil-Werke GmbH Günzburg, Auweg 37
<i>Abel KG., Gebr.</i> Grüna/Sa.	Handschuhfabrik	Wernau/Neckar
<i>Accumulatorenfabrik AG</i> Berlin	Akkumulatoren	Berlin SW 11 Askanischer Platz 3
<i>Achterberg, Georg</i> Berlin	Verlag	Berlin-Lichterfelde Frauenstraße 5
<i>Aechtner & Co.</i> Mülsch	Kleiderstoffe	Reinh. Aechtner Schwarzenbach a. Wald
<i>Actienbrauerei Greussen AG</i> Greussen, Neustadt-Magdeburg, Wittenberge	Brauereien	Wicküler-Küpper-Brauerei AG Wuppertal, Schließfach
<i>Adam & Sohn, C. W.,</i> Staßfurt	Eisen- und Röhrengroßhdlg.	C. W. Adam, Goslar, Am heiligen Graben 3
<i>Adam, Julius Greiner</i> Sonneberg/Th.	Glasspinnerei	Bayreuth Robert-Koch-Straße 11
<i>Ade-Werk</i> Waltershausen	Anhänger- Kupplungen	Offenburg/Baden Englerstraße 9
<i>Aders, Conrad</i> Guben	Leisten- und Rahmenfabrik	Kiel-Wik Wismarer Straße 18

Notes: This table lists the first ten entries reported in the register of West addresses; firm names in italics.

In some cases, it is not possible to identify the geocodes of the place of origin or the place of destination due to ambiguous or missing addresses. For example, there are firms for

which the register only indicates “branches in the Soviet zone” but without specific address. In all such cases, we drop the firms from the data set.

For a significant number of cases, the register only indicates a specific person that can provide information about the re-location of a firm (indicated with “Auskunft”). These cases include firms that were relocated to West Germany after WWII but went out of business until the register was published in 1957. In the empirical analysis, we drop all these cases.

We use the firm name to identify the type of the company. We code all firms as joint-stock companies (AG) if the firm name included the term “AG”, “Aktiengesellschaft”, or “Aktienverein” (or similar spellings). Likewise, we code the firms as limited liability companies if the firm name included the term “GmbH” or “Gesellschaft mbH” (or similar spellings). All other companies are coded as privately-owned companies. These companies include single proprietors and all types of private partnerships with non-limited liability.

We use the industry information to aggregate the firms by industries, relying on the Standard Industry Classification (SIC) (<https://siccode.com/>). Table B2 shows the industry definitions used in the paper and the related industries according to SIC.

Table B2: Industry Classification

Group	Industry Definition	Industries According to SIC
Manufacturing	Foodstuff, Drinks & Tobacco	(20) Food and Kindred Products, (21) Tobacco Products
	Textiles	(22) Textile Mill Products (23) Apparel and other Finished Products
	Wood Products	(24) Lumber and Wood Products (25) Furniture and Fixtures
	Paper	(26) Paper and Allied Products
	Printing & Publishing	(27) Printing, Publishing, and Allied Industries
	Chemicals	(28) Chemicals and Allied Industries
	Petroleum & Coal Products	(29) Petroleum Refining and Related Industries
	Rubber & Plastic	(30) Rubber and Miscellaneous Plastics Products
	Leather	(31) Leather and Leather Products
	Stone, Clay & Glass	(32) Stone, Clay, Glass, and Concrete Products
	Primary Metals	(33) Primary Metal Industries
	Metal Products	(34) Fabricated Metal Products, except Machinery and Transportation Equipment
	Machine Building	(35) Industrial and Commercial Machinery and Computer Equipment
	Electrical Engineering	(36) Electronic and other Electrical Equipment and Components, except Computer Equipment
	Transportation Equipment	(37) Transportation Equipment
	Scientific Instruments	(38) Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks
	Miscellaneous	(39) Miscellaneous Manufacturing Industries
Non-Manufacturing	Agriculture, Forestry & Fishing	(01-09) Agriculture, Forestry, Fishing
	Mining	(10-14) Mining
	Construction	(15-17) Construction
	Transportation	(40-47) Transportation
	Public Utilities	(49) Electric, Gas and Sanitary Services
	Wholesale Trade	(50-51) Wholesale Trade
	Retail Trade	(52-59) Retail Trade
	Finance, Insurance & Real Estate	(60-67) Finance, Insurance, Real Estate
	Services	(70-89) Services

Notes: This table shows the industry definitions used in this paper and the corresponding industries according to SIC using the classification available on <https://siccode.com/>.

B2 Structure of the Data Set

B2.1 Administrative Structure of West Germany

To analyse the determinants of the firms' location choices, we use data at the county level, which means that the alternatives in the discrete choice model are counties. The administrative structure of Germany includes *Landkreise* (rural county) and *Stadtkreise* (city county). Larger cities were often administered as a separate city county, while the surrounding parishes were part of the rural county. Over time, some parishes were incorporated in the city counties. Therefore, to match pre-war with post-war data, we merged all city with rural counties. In total, the data includes 467 counties (including one county for West Berlin). The data excludes all counties in the Saarland since the Saarland was separated from West Germany until 1957.

B2.2 Border Sub-Samples

In Table 7, we test the determinants of firm choice for firms that were, in the pre-war period, located close to the post-war East-West border. To create the border sub-samples, we assign all firms to grid cells of 0.5 x 0.5 degrees, based on the pre-war location in East Germany. In the first sub-sample, we use a narrow definition and include only firms that were in a band of grid cells that touched or crossed the border. In the second sub-sample, we include all firms from the first sub-sample as well as all firms that were originally located in all adjoining grid cells.

B2.3 Places Temporary Occupied by US and British troops

For each company, we identify whether the old place in East Germany was in an area temporarily occupied by the Western Allies (US and British troops) in April and May 1945. The coding is based on the map "Central Europe, 1944. The End of War. Final Operations, 19 April – 7 May 1945" ([Link](#)). For some smaller villages along the Eastern frontline border, we used additional information from Wikipedia and other online sources to check whether there is information about temporary US or British occupation in 1945.

B3 Distance

$\ln(\text{Distance})$ is the natural logarithm of the geodesic distance (in km) between a firm's old location in East Germany and the geographic centroid of the potential county of destination in West Germany.

B4 Agglomeration Proxies

We use (a) population density, (b) firm revenues per capita, and (c) Domestic Market Potential (DMP) as proxies for agglomeration effects.

B4.1 Population Density

The variable $\ln(\text{Pop. Density } 1933)$ [$\ln(\text{Pop. Density } 1950)$] indicates the natural logarithm of the population density in a county (inhabitants per km²) based on data from the [West] German population census of 1933 [1950]. The 1933 [1950] census data are published in Statistisches Reichsamt (1936) [Statistisches Bundesamt (1952)]. Data on the area size of a county is from the 1950 population census.

B4.2 Population Density

The variable $\ln(\text{Revenue pc } 1935)$ [$\ln(\text{Revenue pc } 1950)$] is the natural logarithm of the revenues of firms located in a county in 1935 [1950] divided by the population in 1933 [1950]. We divide 1935 revenue data by population from 1933 since there is no county-level population data for 1935 available. Data on firm revenues in 1935 [1950] is from the sales-tax statistic (*Umsatzsteuerstatistik*), which is published in Statistisches Reichsamt (1939) [Statistisches Bundesamt (1955)], data on the county population is from Statistisches Reichsamt (1936) [Statistisches Bundesamt (1952)].

B4.3 Domestic Market Potential

The variable $\ln(\text{DMP } 1933)$ is the natural logarithm of the Domestic Market Potential (DMP), which is defined as

$$\text{DMP}_c = \sum_{j \neq c} \frac{\text{Pop}_j^{1933}}{(D_{jc})^2}$$

with Pop_j^{1933} (the county population of county j based on the 1933 census data), which we divide with the squared distance between county j and c . We use the squared distance in the denominator to give more distant counties a lower weight. Then, we compute the DMP for each county c by summing up all distance-weighted county populations. For 1950, we compute $\ln(\text{DMP } 1950)$ in the same way by using data from the 1950 population census. Population data for 1933 [1950] is from Statistisches Reichsamt (1936) [Statistisches Bundesamt (1952)].

Since West-Berlin was an exclave and trade restricted, given that it had to pass East German territory, we set the value for DMP to 0 for the county West-Berlin and we set the population of West-Berlin to 0, when computing DMP for all other counties for 1933 and 1950, respectively. Note that we also control for West Berlin in the regressions with all controls.

B5 Control Variables

B5.1 Occupation Zones

Berlin is a dummy variable that is one for West Berlin, and zero for all other counties. *British Occupation* [*French Occupation*] is a dummy variable that is one for all counties that were part of the British [French] occupation zone, and zero for all other counties. The reference group are counties in the US occupation zone.

B5.2 Universities

University 1950 is a dummy variable that is one if a university was in the respective county in 1950, and zero for all counties without university. $\ln(\text{Distance to University 1950})$ is the natural logarithm of the minimum distance between a university in 1950 and the centroid of a county in our data set.

B5.3 Sea Ports

The variable *Sea Port* is equal to one for the counties where the three major North-Sea ports were located in West Germany (Hamburg, Bremen (including Bremerhaven), and Emden), and zero for all other counties.

B5.4 Highway

The variable *Highway* is equal to one if the county had access to the highway system in 1950, and zero for all other counties. We define access to the highway system if at least one highway crossed a county or parts of it. Information on the location of highways is from the map “ADAC Autobahn-Karte” published in 1950 by the *Allgemeiner Deutscher Automobil-Club* (ADAC), the main German automobile association. We geocoded the map and matched it with our county-level data. The map is accessible via the David Rumsey Map Collection ([Link](#)).

B5.5 Latitude and Longitude

The variable *Latitude* [*Longitude*] indicates the centroid latitude [longitude] of a county.

B5.6 Eastern Border

The variable *Eastern Border* is equal to one for all West German counties that were at the Eastern border (including the border to East Germany and Czechoslovakia). These counties formed the so-called Zonal Border Area [*Zonenrandgebiet*], which later received financial subsidies.

B5.7 Tax Rate

The variable *Tax Rate* indicates an estimate for the average *Gewerbesteuer* in a county. The German *Gewerbesteuer* is a business tax that is structured such that the community can define a “tax factor”, the so-called *Hebesatz*, which is then applied on the tax assessment basis, resulting in different local tax burdens. For example, a tax factor of 350 percent means that the tax assessment basis is multiplied by 3.5 before charging the tax. In our period of observation, the *Gewerbesteuer* could be levied either on profits and capital or on the wage total. Most communities levied the tax only on profits and capital, but there were also differences across states. In the state of Nordrhein-Westfalen, the communities levied the tax on a combination on both, and, in Hessen and Rheinland-Pfalz, the tax on wage total was levied only if the tax burden was larger than when levied on profits and capital.

Given the typical taxation policies of the communities at the time, we assume that a one-percentage-point tax factor on profits and capital equals a 1/16-percentage-points tax factor on the wage total. Using this assumption, we computed a unique measure for the tax factor over all communities for which we have data. If there is only one community listed per county, we take this community tax factor as an estimate for the average tax factor in this county. If there is information on more communities, than we take the unweighted mean tax factor within a county. Note that the communities for which information on the tax factor is available are typically the largest cities in the respective county. To construct the data, we use information from Deutscher Städtetag (1951, p. 367-374), which lists the tax factor for 473 communities. This data allows us to construct the average tax factor for 283 counties. Data are missing in particular for very rural counties, but these counties did not account for a large fraction of relocated firms so that the potential bias that may be created by the sample restriction should be small. To facilitate the interpretation of the coefficients, we divided the estimated tax factors by 100, which means that an increase of the variable *Tax Rate* by one unit implies an increase of the tax factor by 100 percentage points.

B5.8 Expellees

The variable *Expellees 1950 %* indicates the percentage share of the expelled population relative to the total population of a county. Data is from Braun and Franke (2021). We adjusted this data to match it with our data at the county level.

B6 Industry Shares

In section 6.5, we test whether the location choice was affected by the pre-war industry structure. To measure the differences in the pre-war industry structure, we use data from the

1925 employment census (Deutsches Reich 1927), which provides information about the number of employees by counties and industries. The crude data was collected by Felix Kersting, who thankfully shared his data with us. *Industry Share %* is the total number of employees in a specific industry and county relative to the total number of employees in this county.

The employment census distinguishes between 27 different employment groups, including the common industries as well as several categories for employment in the public sector. To match the employment shares with our firm data, we combined several sectors. Figures 8 and 9 distinguish between the following industries (1925 classification in brackets, classification number in parentheses): “Textiles” [textile industry (X) and clothing (XVII)]; “Machine-building & Transportation” [construction of machinery, apparatus, and vehicles (VII)]; “Foodstuff & Tobacco” [food, drinks and tobacco (XVI)]; “Paper & Printing” [paper and printing (XI)]; “Chemicals” [chemical industry (IX)]; “Metal Products” [products of iron, steel, and metal (VI)]; “Wood Products” [wood products (XIV)]; “Electrical Engineering & Instruments” [electrical engineering and precision engineering (VIII)]; “Construction” [construction industry (XVIII)].

We do not run separate regressions for very small industries with less than 100 firms and we do also not run separate regressions for the large service sector industries (see Figure 7) since the census data uses only rather broad definitions. These broad definitions do not allow for a meaningful analysis, since we cannot distinguish between different parts of the service sector (e.g., data is only available for “*Handelsgewerbe*” (commercial enterprises, XX), which includes retail trade, wholesale trade, some financial services, and real estate businesses).

Data on employment is available on the sub-county level, indicating employment of rural parts of counties and employment in larger cities separately. Due to administrative reforms in the 1930, we had to adjust the data from 1925 such that we can match it with our data set, for example by merging several counties. Note that there are two special cases:

In 1942, the county Salzgitter was formed by some communities that belonged to two different counties, Goslar and Wolfenbüttel. Since we cannot identify employment numbers for these communities, we take the mean employment share of the counties of Goslar and Wolfenbüttel for the newly established county of Salzgitter.

The county Bad Reichenhall (Bavaria) was originally part of the county Berchtesgaden. Since there are no separate employment figures for Bad Reichenhall in 1925, we take the same employment shares as for Berchtesgaden.

References for the Appendix

- Braun, Sebastian Till and Richard Franke (2021). A County-Level Database on Expellees in West Germany, 1939-1961. *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte* 108(4), 522-540.
- Deutsches Reich (1927). *Berufszählung 1925: Die berufliche und soziale Gliederung der Bevölkerung des deutschen Reichs*. Verlag für Sozialpolitik, Wirtschaft und Statistik, Berlin.
- Deutscher Städtetag (1951). *Statistisches Jahrbuch Deutscher Gemeinden*. 39. Jahrgang. Alfons Bürger Verlag.
- Interessengemeinschaft der in der Ostzone Enteigneten Betriebe (1957). *Westadressen-Verzeichnis der Industrie- und Handelsbetriebe, Banken und Versicherungen aus der sowjetisch besetzten Zone und dem Ostsektor Berlins*. Verlag Klaus Edgar Herfurth, Frankfurt/Main.
- Statistisches Bundesamt (1955). *Die Umsätze der Umsatzsteuerpflichtigen und deren Besteuerung (Ergebnisse der Statistik über die Umsatzsteuerveranlagung für 1950)*. Kohlhammer, Stuttgart.
- Statistisches Bundesamt (1952). *Statistisches Jahrbuch für die Bundesrepublik Deutschland*. Kohlhammer, Stuttgart.
- Statistisches Reichsamt (1939). *Umsatzsteuerstatistik 1935. IV. Teil. Ergebnisse der Verwaltungsbezirke*. Verlag für Sozialpolitik, Wirtschaft und Statistik, Berlin.
- Statistisches Reichsamt (1936). *Amtliches Gemeindeverzeichnis für das Deutsche Reich auf Grund der Volkszählung 1939*. Verlag für Sozialpolitik, Wirtschaft und Statistik, Berlin.