Location Fundamentals, Agglomeration Economies, and the Geography of Multinational Firms

HARVARD | BUSINESS | SCHOOL

Location Fundamentals, Agglomeration Economies, and the Geography of Multinational Firms

Laura Alfaro
Maggie Xiaoyang Chen

Working Paper 17-014

Location Fundamentals, Agglomeration Economies, and the Geography of Multinational Firms

Laura Alfaro
Harvard Business School

Maggie Xiaoyang Chen
George Washington University

Working Paper 17-014

Location Fundamentals, Agglomeration Economies, and the Geography of Multinational Firms*

Laura Alfaro†
Harvard Business School and NBER

Maggie Xiaoyang Chen‡
George Washington University

August 2016

Abstract

Multinationals exhibit distinct agglomeration patterns which have transformed the global landscape of industrial production (Alfaro and Chen, 2014). Using a unique worldwide plant-level dataset that reports detailed location, ownership, and operation information for plants in over 100 countries, we construct a spatially continuous index of pairwise-industry agglomeration and investigate the patterns and determinants underlying the global economic geography of multinational firms. In particular, we run a horse-race between two distinct economic forces: location fundamentals and agglomeration economies. We find that location fundamentals including market access and comparative advantage and agglomeration economies including capital-good market externality and technology diffusion play a particularly important role in multinationals' economic geography. These findings remain robust when we use alternative measures of trade costs, address potential reverse causality, and explore regional patterns.

JEL codes: F2, D2, R1

Key words: multinational firm, economic geography, agglomeration, location fundamentals, agglomeration economies

---

*We thank participants in the Workshop on Structural Transformation and Jim Anderson, Bruce Blonigen, Gilles Duranton, James Harrigan, Keith Head, Tarun Khanna, Jim Markusen, Keith Maskus, Mike Moore, Henry Overman, John Ries, Roberto Samaneigo, and Tony Yezer for valuable comments and suggestions in various stages of the project and William Kerr for kindly providing us the patent concordance data. Hayley Pallan, Elizabeth Meyer, and Hillary White provided superb research assistance.

†Email: lalfaro@hbs.edu; Phone: 617-495-7981; Harvard Business School, Morgan 293, Boston, MA 02163.

‡Email: xchen@gwu.edu; Phone: 202-994-0192; Department of Economics, George Washington University, 2115 G ST, NW, #367, Washington, DC 20052.

1 Introduction

An exponential increase in flows of goods, capital, and ideas is one of the most prominent economic trends in recent decades. A key driver of this phenomenon is cross-border production, investment, and innovation led by multinational corporations (MNCs). Multinational affiliate sales as a share of world GDP have more than doubled in the past two decades, increasing from close to 25 percent in 1990 to more than 50 percent in 2014.[1] This explosion of MNC activities is rapidly transforming the global landscape of industrial production, precipitating the emergence of new industrial clusters around the world. Firms that agglomerated in, for example, Silicon Valley and Detroit now have subsidiary plants clustering in Bangalore and Slovakia (termed, respectively, the Silicon Valley of India and Detroit of the East). Understanding the agglomeration, and more broadly the economic geography of MNCs and their role in shaping the global industrial landscape as a result has become increasingly important for shaping policies and promoting the benefits of globalization.

Recent evidence shows that MNCs exhibit distinct agglomeration patterns (Alfaro and Chen, 2014). In contrast to domestic production which emphasizes domestic geography and natural advantage, multinational production stresses foreign market access and international comparative advantage. Moreover, as highlighted in a growing literature (e.g., Helpman, Melitz, and Yeaple, 2004; Antras and Helpman, 2004, 2008; Yeaple, 2009; Chen and Moore, 2010), the economic attributes and organization of multinationals are, by selection, distinctively different from average domestic firms. Their vertically integrated organization, greater productivity, and higher capital- and knowledge-intensities relative to domestic firms all suggest that MNCs are likely to exhibit different agglomeration motives.

This chapter seeks to grasp the complexity in the various drivers of MNC location decisions. In particular, we run a horse-race between two distinct economic forces: location fundamentals (also referred to as "first nature") of multinational production (MP) and agglomeration economies (also known as "second nature"). The location fundamentals of MP, as stressed in the international trade literature, consist of primarily foreign market access (multinationals choose to produce in large foreign markets to avoid trade costs) and comparative advantage (multinationals locate production in countries with desired factor abundance and low factor prices).[2] In contrast, agglomeration economies, the study of which dates back to Marshall (1890), stress the benefits of geographic proximity between firms, including lower transport costs between input suppliers and final good producers, labor and capital-good market externalities, and technology diffusion. While existing studies offer separate evidence on the roles of the above two categories of factors in multinationals' location decisions, how these factors jointly influence multinationals' global

[1] Source: UNCTAD, World Investment Report (2015).

[2] While comparative advantage is defined here in the context of neoclassical trade theory, other country factors such as institutional characteristics and physical locations can also play a role in firms' location decisions and are sometimes considered as part of comparative advantage (see, e.g., Nunn, 2007). As described in Section 4.1, our empirical specification controls for all host-country specific factors when constructing the location fundamentals of multinational production measure.

economic geography given their organizational structure and capital- and knowledge-intensive production requires further exploration.

An evaluation of the patterns and causes of MNC economic geography faces, however, several key challenges. First, the measurement of agglomeration has been a central issue in the economic geography literature. Traditional indices that define agglomeration as the amount of activity located in a particular geographic unit omit agglomerative activities separated by administrative and geographic borders and can be affected by the extent of industrial concentration. Second, distinguishing between the effects of MP's location fundamentals and agglomeration economies is complicated by the difficulty of measuring the factors quantitatively. Further, their common propensity to lead MNCs to locate next to each other makes it difficult to separate their relative effects. Third, identifying the causal effects of location fundamentals and agglomeration economies is a key challenge in empirical analyses of economic geography. Both types of factors can endogenously reflect the patterns of MNC agglomeration. Finally, quantifying the global patterns of MNC economic geography requires cross-country data that document multinational production at the plant, instead of firm, level.

To overcome the above challenges, our empirical analysis proceeds in the following steps. First, we quantify the global agglomeration of MNCs using WorldBase, a worldwide plant-level dataset that provides detailed location, ownership, and activity information for establishments in more than 100 countries. Its broad cross-country coverage enables us to depict worldwide patterns of MNC economic geography. Moreover, the dataset's detailed location and operation information for over 43 million plants, including multinational and domestic, offshore and headquarters establishments, makes it possible to compare the geographic patterns of different types of establishments. The physical location of each establishment allows us to construct indices of agglomeration using precise latitude and longitude codes for each plant and the distance and trade cost between each pair of establishments.

Second, we construct a spatially continuous index of agglomeration for pairwise industries (also referred to as coagglomeration).[3] We obtain latitude and longitude codes for each establishment in the data based on plant-level physical location information and compute not only the distance but also the trade cost that accounts for other forms of trade barriers between each pair of establishments. Following an empirical methodology introduced by Duranton and Overman (2005) (henceforth, DO) and extended in Alfaro and Chen (2014), we then employ a Monte-Carlo approach that compares the actual geographic density of plants in each industry pair with counterfactual densities. This procedure separates agglomeration from the general geographic concentration of multinationals and deals with the effect of industrial concentration. Industry pairs that exhibit greater geographic density than the counterfactuals are considered to exhibit significant evidence of agglomeration.

[3] We use the term "agglomeration" broadly to refer to both within- and between-industry agglomeration (the latter sometimes referred to as "coagglomeration"). The broad usage of the term "agglomeration" is fairly common in the literature.

As in Alfaro and Chen (2014), we construct the agglomeration index for each pairwise industry to help disentangle the effects of location fundamentals and various agglomeration forces. As noted by Ellison, Glaeser, and Kerr (2010) (henceforth, EGK), while location fundamentals and agglomeration economies tend to predict agglomeration among firms in the same industry, their predictions of which industry pairs should agglomerate vary significantly. Compared to firms in the same industries, firms from different industry pairs often exhibit greater variation in their relatedness in production, factor markets, and technology space, thereby displaying different agglomeration incentives.[4] Exploring the pairwise-industry agglomeration of MNCs thus makes it possible to separate the effects of location fundamentals and the various agglomeration economies.

Third, after computing the actual agglomeration index of MNCs, we construct an expected index of MNC agglomeration to capture the effect of location fundamentals. This index reflects the geographic distribution of MNC plants predicted exclusively by the location fundamentals of multinational production including, among others, foreign market size, trade costs, and comparative advantage. Specifically, we invoke a two-step procedure. In the first step, we estimate a conventional gravity-type MP equation and examine the effects of market access and comparative advantage as well as other location factors in multinationals' location decisions. Based on the estimates, we obtain the location patterns of MNC plants predicted by the location fundamentals and, in the second step, construct an index of agglomeration using the predicted, instead of actual, locations. This index represents the expected degree of pairwise-industry agglomeration based on industry pairs' common location fundamentals.

Fourth, controlling for the agglomeration predicted by location fundamentals and all industry specific factors, we examine the degree to which proxies of agglomeration forces, including between-industry input-output linkages, labor demand similarity, technology spillover, and a new measure of capital-good market externality, explain the variations in the agglomeration index of multinational firms. We construct the proxies of agglomeration forces using lagged, disaggregated U.S. industry account data to mitigate the potential reverse causality concern, as it is not very likely that U.S. industries' production, factor, and technology linkages are a result of worldwide MNCs' agglomeration patterns.

To further alleviate concerns of endogenous agglomeration economy measures, we examine regional agglomeration patterns from which the United States is excluded. If U.S. domestic industry-pair relationships could be affected by the agglomeration of MNCs in the U.S., then one would expect that the former would not be affected by the agglomeration of MNCs located in other regions, like Europe.[5]

Our analysis presents a rich array of new findings that shed light on the global agglomera-

[4]For example, firms in the automobile industry may agglomerate because of both location fundamentals and any of the agglomeration economies whereas firms in the automobile and steel industries are likely to agglomerate mainly because of their production linkages.

[5]In Alfaro and Chen (2014), we investigate the process of agglomeration. Exploring the dynamics in MNCs' offshore agglomeration sheds light on the formation of MNC clusters and mitigates the possibility of reverse causation between our measures of location fundamentals, agglomeration economies, and MNCs' agglomeration patterns.

tion of MNCs. First, the location fundamentals of multinational production, although playing a significant and vital role, are not the only driving forces in the patterns of MNC offshore agglomeration. As shown in Alfaro and Chen (2014), agglomeration economies, especially capital-good market externality and technology diffusion, are crucial determinants of MNCs' overseas location decisions. When comparing the relative importance of location fundamentals and agglomeration economies, however, we find the effect of location fundamentals to exceed the cumulative impact of agglomeration forces. A one-standard-deviation increase in the former is associated with a 0.31 standard-deviation increase in the extent of MNCs' offshore agglomeration at the 200 km level, whereas the cumulative effect of agglomeration economies is around 0.17.

Second, as suggested by the agglomeration patterns, the relative importance of location fundamentals and agglomeration economies varies significantly between MNC offshore and domestic plant agglomeration and between MNC offshore and headquarters agglomeration. Comparing the agglomeration of MNC offshore and domestically owned plants, we find MNC plants, reflecting their high capital- and innovation-intensities, to be significantly more influenced by capital-good market and technological agglomeration factors. The under-provision of capital goods in many host countries increases MNCs' incentives to locate in proximity to one another overseas and take advantage of agglomeration economies. Moreover, location fundamentals and capital-good market externality exert a stronger effect on the offshore agglomeration of MNC subsidiary establishments, while technology diffusion and labor market externalities are the leading forces behind the agglomeration of headquarters. Vertical production linkages, in contrast, matter for offshore clustering only.

These results are consistent with the increasing segmentation of activities within the boundary of multinational firms, in particular, the market-seeking and the input-sourcing focuses of offshore production and the emphasis of headquarters on knowledge intensive activities such as R&D, management, and services. The findings also remain largely robust when examining regional agglomeration patterns, and restricting the analysis to Europe which yield additional insights.

The rest of the chapter is organized as follows. Section 2 reviews the related literature. Section 3 discusses the methodology used in this chapter to construct pairwise-industry agglomeration indices. Section 4 describes the methodology used to measure location fundamentals and agglomeration economies. Section 5 describes the cross-country establishment data. Section 6 reports the stylized facts and econometric evidence on the determinants of MNC economic geography. Section 7 presents additional analysis that examines agglomeration in Europe. The last section concludes.

2 Related Literature

Our work builds on three broad strands of literature. First, an extensive literature in international trade has shed important theoretical and empirical light on the role of location fundamentals in MNCs' decisions to invest abroad. Two main motives of foreign investment have been stressed by studies in this literature. First, firms may choose to produce overseas to avoid trade costs. This strategy, referred to as the market access (or tariff jumping) motive, leads firms to duplicate production processes in countries (see, e.g., Markusen, 1984 and Markusen and Venables, 2000). Second, firms may choose to locate different stages of production in countries where the factor used intensively is abundant. This strategy is referred to as the comparative advantage motive (see, e.g., Helpman, 1984). These two motives, leading to horizontal and vertical FDI respectively, have been synchronized in the knowledge-capital model developed by Markusen and Venables (1998) and Markusen (2002) and examined in a number of empirical studies.[6] This strand of literature provides a theoretical and empirical foundation for the location fundamental portion of our research.

A second related literature consists of the extensive body of research in regional and urban economics that has been devoted to evaluating the importance of Marshallian agglomeration forces in domestic economic geography.[7] Marshall (1890) first introduced the idea that concentrations of economic factors, such as knowledge, labor, and inputs, can generate positive externalities. Data restrictions have impeded the progress of studying economic geography at a global scale; most related research has focused on a geographic area such as the United States (Rosenthal and Strange, 2001) or the United Kingdom (Overman and Puga, 2009), further discussed below. As in Alfaro and Chen (2014), the Worldbase database allows us to explore the global economic geography of MNCs in this chapter.

As noted earlier, a central issue in studies of agglomeration concerns the measurement of agglomeration. In an influential paper, Ellison and Glaeser (1997) introduce a "dartboard" approach to construct an index of spatial concentration. The index compares the observed distribution of economic activity in an industry to a null hypothesis of random location and controls for the effect of industrial concentration, an issue noted to affect the accuracy of previous indices. Using Ellison and Glaeser's (1997) index to evaluate the importance of agglomeration forces in explaining the localization of U.S. industries, Rosenthal and Strange (2001) find both labor market pooling and input-output linkages to have a positive impact on agglomeration. Also employing Ellison and Glaeser's (1997) index, Overman and Puga (2009) examine the role of labor market pooling and input sharing in determining the spatial concentration of UK manufacturing establishments. They find sectors whose establishments experience more idiosyncratic employment volatility and that use localized intermediate inputs are more spatially concentrated.

The study by DO advances the literature by developing a spatially continuous concentration index that is independent of the level of geographic disaggregation. Applying this index, EGK employ an innovative empirical approach that exploits the coagglomeration of U.S. industries to disentangle the effects of Marshallian agglomeration economies. They find, as in Rosenthal and Strange (2001), a particularly important role for input-output relationships.

[6] The analysis by Carr, Markusen, and Maskus (2001), Yeaple (2003a), and Alfaro and Charlton (2009), for example, offers empirical support for both types of motives.

[7] See Ottaviano and Puga (1998), Head and Mayer (2004), Ottaviano and Thisse (2004), Rosenthal and Strange (2004), Duranton and Puga (2004), Puga (2010), and Redding (2010, 2011) for excellent reviews of these literatures.

Exploring the role of agglomeration economies in MNCs' location patterns also relates the present chapter to a third strand of literature in international trade that emphasizes the advantage of proximity between customers and suppliers. Several studies (see, e.g., Head, Ries, and Swenson, 1995; Bobonis and Shatz, 2007) show that MNCs with vertical production linkages tend to agglomerate regionally within a country. Alfaro and Chen (2014), using also the D&B data, construct a spatially continuous index of agglomeration extending the DO methodology and analyze the different patterns underlying the global economic geography of multinational and non-multinational firms. We uncover new stylized facts that suggest the offshore clusters of multinationals are not a simple reflection of domestic industrial clusters. We find that agglomeration economies including technology diffusion and capital-good market externality play a more important role in the offshore agglomeration of multinationals than the agglomeration of domestic firms. These findings remain robust when exploring the process of agglomeration.

Our current analysis, contributes to the above literature and extends it in several important ways. Instead of focusing on domestic agglomeration patterns in industrialized countries like the U.S., our analysis offers a perspective on the structure of industrial agglomeration at the world and region level. In particular, we investigate how the most mobile and distinctive group of firms—multinationals—agglomerate domestically and overseas. Importantly, we incorporate the location fundamentals of MNCs into the analysis of agglomeration and develop a new quantitative measure to quantify the role of location fundamentals in MNCs' spatial concentrations. Further, we evaluate how agglomeration economies, particularly the value of external scale economy in capital goods and knowledge, affect MNCs relative to domestic firms, given MNCs' vertically integrated organizational form and high demand for capital goods and technologies. Our results show that location fundamentals matter and that capital-good externality and technology diffusion, factors that have not been emphasized in this literature, exert an important effect on the agglomeration of MNCs.

3 Quantifying Agglomeration: Methodology

In this section, we describe the empirical methodology used to quantify the global agglomeration of multinational firms. As noted in Head and Mayer (2004), measurement of agglomeration is a central challenge in the economic geography literature.[8] Continuous effort has been devoted to designing an index that accurately reflects the agglomeration of economic activities. One of the latest progresses in this literature is DO.

DO construct this index to measure the significance of same-industry agglomeration in the U.K. The index has then been adapted by EGK to investigate the coagglomeration of U.S.

[8] More recently, Duranton and Kerr (2015) have also noted the difficulty of obtaining appropriate data in order to measure agglomeration. They emphasize the importance and potential contribution that new data will allow for research into agglomeration-related issues. In line with our research, they suggest the need for further research into the role of multinationals and their subsidiaries in agglomeration economies given new availability of data sources.

industries. We extend this index to a global context to measure the degree of coagglomeration of multinational firms worldwide. Because it accounts for continuity in space, the index is well suited for cross-country studies as shown in Alfaro and Chen (2014). In this chapter, we also expand the original index's focus on distance as the main form of trade cost to a measure that accounts for various forms of trade costs (distance, tariffs, etc.).

We first describe the empirical procedure and then discuss its main advantages and shortcomings. The empirical procedure to construct the index involves three steps. To compare global location patterns of MNC subsidiaries, headquarters, and domestic firms we repeat the procedure for each type of establishment.

Step 1: Kernel estimator We first estimate an actual geographic density function for each pair of industries. Note that even when the locations of nearly all establishments are known with a high degree of precision (such as in the data we use, as described in section 5), distance (as well as estimated trade cost) is an approximation of the true trade cost between establishments. One source of systematic error, for example, is that travel time for any given distance might differ between low- and high-density areas. Given the potential noise in the measurement of trade costs, we follow DO in adopting kernel smoothing when estimating the distribution function.

Let τ_{ij} denote the distance between establishment i and j. For each industry pair k and $\widetilde{k}$, we obtain a kernel estimator at any point τ (e.g., $K_{k\widetilde{k}}(\tau)$):

$$f_{k\widetilde{k}}(\tau) = \frac{1}{n_k n_{\widetilde{k}} h} \sum_{i=1}^{n_k} \sum_{j=1}^{n_{\widetilde{k}}} K\left(\frac{\tau - \tau_{ij}}{h}\right), \tag{1}$$

where n_k and $n_{\widetilde{k}}$ are the number of plants in industries k and $\widetilde{k}$, respectively, h is the bandwidth, and K is the kernel function. We use Gaussian kernels with the data reflected around zero and the bandwidth set to minimize the mean integrated squared error.[9] This step generates a kernel estimator for each of the $7,875$ $(= 126 \times 125/2)$ manufacturing industry pairs in our data.[10]

In addition to estimating the geographic distribution of establishment pairs, we can also treat each worker as the unit of observation and measure the level of agglomeration among workers. To proceed, we obtain a weighted kernel estimator by weighing each establishment by employment size, given by

$$f^{w}_{k\widetilde{k}}(\tau) = \frac{1}{h\sum_{i=1}^{n_k}\sum_{j=1}^{n_{\widetilde{k}}}(r_i r_j)} \sum_{i=1}^{n_k}\sum_{j=1}^{n_{\widetilde{k}}} r_i r_j K\left(\frac{\tau - \tau_{ij}}{h}\right) \tag{2}$$

[9]Although we follow DO and EGK in obtaining kernel estimators, a less computationally intensive approach that yields similar properties would be to look at cumulative distances.

[10]Identical industry pairs (126 observations) are dropped from the analysis because, as explained earlier, we rely on industry-pair variations in relatedness in production, factor demand, and technology to disentangle the effects of location fundamentals and various agglomeration economies. Identical industry pairs exhibit all dimensions of relatedness and lack the needed variation. Moreover, as we explain in Section 4, the measures of location fundamentals and agglomeration economies used in this chapter, by design, capture only between-industry relationships. The main empirical analysis is performed at the SIC 3-digit level. This level of industry disaggregation is dominated by the availability of control variables, as described in Section 4.

where r_i and r_j represent, respectively, the number of employees in establishments i and j. We do this for each of the $7,875$ industry pairs.

Step 2: Counterfactuals and global confidence bands To obtain counterfactual estimators, we estimate the geographic distribution of the manufacturing multinationals as a whole in order to control for factors that affect all manufacturing multinational plants. We proceed by drawing, for each of the 7,875 industry pairs, 1,000 random samples, each of which includes two counterfactual industries. Given our goal of obtaining, in this step, the overall agglomeration patterns of MNCs, the random samples are drawn from the entire set of MNC establishment locations.[11] Note that to control for the potential effect of industry concentration, it is important that the counterfactual industry in each sample has the same number of observations as the actual data. We then calculate the bilateral distance between each pair of establishments and obtain a kernel estimator, unweighted or weighted by employment, for each of the 7,875,000 samples. This gives $1,000$ kernel estimators for each of the $7,875$ industry pairs.

We compare the actual and counterfactual kernel estimators at various distance thresholds, including 200, 400, 800, and 1,600 kilometers (the maximum threshold being roughly the distance between Detroit and Dallas and between London and Lisbon). We compute the 95% global confidence band for each threshold distance. Following DO, we choose identical local confidence intervals at all levels of distance such that the global confidence level is 5%. We use $\overline{f}_{k\widetilde{k}}(\tau)$ to denote the upper global confidence band of industry pair k and $\widetilde{k}$. When $f_{k\widetilde{k}}(\tau) > \overline{f}_{k\widetilde{k}}(\tau)$ for at least one $\tau \in [0, T]$, the industry pair is considered to agglomerate at T and exhibit greater agglomeration than counterfactuals. Graphically, it is detected when the kernel estimates of the industry pair lie above its upper global confidence band.

Step 3: Agglomeration index We now construct the agglomeration index. For each industry pair k and $\widetilde{k}$, we obtain

$$agglomeration_{k\widetilde{k}}(T) \equiv \sum\nolimits_{\tau=0}^{T} \max\left(f_{k\widetilde{k}}(\tau) - \overline{f}_{k\widetilde{k}}(\tau), 0\right) \tag{3}$$

or employment-weighted

$$agglomeration^{w}_{k\widetilde{k}}(T) \equiv \sum\nolimits_{\tau=0}^{T} \max\left(f^{w}_{k\widetilde{k}}(\tau) - \overline{f}^{w}_{k\widetilde{k}}(\tau), 0\right). \tag{4}$$

The index measures the extent to which establishments in industries k and $\widetilde{k}$ agglomerate at threshold distance T and the statistical significance thereof. When the index is positive, the level of agglomeration between industries k and $\widetilde{k}$ is significantly greater than that of the counterfactuals.

[11] An alternative approach would be to use all existing, including domestic and MNC, establishment locations as the counterfactuals. This would help to control for the effect of general location factors instead of those that affect primarily MNCs' location decisions. In Section 6.3, we perform an analysis in that direction by employing domestic establishments as the benchmark and comparing the agglomeration patterns of MNC and domestic plants.

There are two requirements for the construction of the DO index. First, availability of physical location information for each establishment at the most detailed level. The WorldBase dataset, described in section 5, supplemented by a geocoding software, satisfies this requirement. Second, the empirical procedure adopted to construct the index uses a simulation approach that is computationally intensive, especially for cross-country studies and large datasets.

4 Measuring Location Fundamentals and Agglomeration Economies

We now turn to economic factors that could systematically account for the observed agglomeration patterns of MNCs. Incorporating the multinational firm theory with the literature of economic geography, the location decisions of multinational firms can be viewed as a function of two categories of factors. One consists of location fundamentals of MP that motivate MNCs to invest in a given country including market access and comparative advantage. The other is agglomeration economies, which includes vertical production linkages, externality in labor markets and capital-good markets, and technology diffusion.

4.1 MP Location Fundamentals

To quantify MP location fundamentals we construct a measure that incorporates an empirical approach from the multinational firm literature with the agglomeration index methodology and invoking a two-step procedure.

Step 1 In the first step, we seek to obtain estimates of multinational actitivity predicted by location fundamentals including market size, trade cost, comparative advantage and natural advantage, among other related characteristics. To obtain such estimates, we consider two alternative specifications.

In the first specification, we estimate a conventional empirical equation following Carr, Markusen and Maskus (2001), Yeaple (2003a), and Alfaro and Charlton (2009). Using a conventional empirical specification enables us to assess how MP location fundamentals commonly stressed by previous studies affect MNCs' agglomeration patterns. Specifically, we consider the following specification:

$$\begin{aligned} y_{c\widetilde{c}k} = {} & \gamma_0 + \gamma_1 marketsize\_size_{c\widetilde{c}} + \gamma_2 distance_{c\widetilde{c}} + \gamma_3 skill\_diff_{c\widetilde{c}} \\ & + \gamma_4 skill\_diff_{c\widetilde{c}} \times skillintensity_k + \gamma_5 tariff_{c\widetilde{c}k} + \gamma_6 tariff_{\widetilde{c}ck} + \mu_{ck} + \mu'_{\widetilde{c}k} + \varepsilon_{c\widetilde{c}k} \end{aligned} \qquad (5)$$

where $y_{c\widetilde{c}k}$ denotes either the number or the total employment of subsidiaries in country $\widetilde{c}$ and industry k owned by MNCs in country c, $marketsize\_ave_{c\widetilde{c}}$ is average market size proxied by the GDP of home and host countries,[12] $distance_{c\widetilde{c}}$ is the distance, $skill\_diff_{c\widetilde{c}}$ represents the

[12] We consider, in addition to GDP, market potential which is the sum of domestic and distance-weighted export market size of the home and host countries.

difference in skill endowment, measured by average years of schooling, between the home and the host countries (i.e., $skill_{\widetilde{c}} - skill_c$), $skillintensity_k$ is the skilled labor intensity proxied by share of non-production workers for each industry, $tariff_{c\widetilde{c}k}$ and $tariff_{\widetilde{c}ck}$ are the levels of tariffs set by the host country $\widetilde{c}$ on the home country c and vice versa in industry k, and $\varepsilon_{c\widetilde{c}k}$ are the residuals. In addition to the above variables, host-country characteristics such as institutional and physical infrastructure could also affect multinationals' location decisions.[13] We thus include vectors of country-industry dummies, μ_{ck} and $\mu'_{\widetilde{c}k}$, to control for all country-industry specific factors such as institutional quality, physical infrastructure, domestic industry size, and economic policies.[14]

We estimate equation (5) using Poisson quasi-MLE (QMLE).[15] If market access is a significant motive in MNCs' investment decisions, we expect the effects of market size and trade cost (measured by distance and tariffs) to be positive, that is, $\gamma_1 > 0$, $\gamma_2 > 0$, and $\gamma_5 > 0$. If comparative advantage is a significant motive, we expect the effect of trade cost to be negative and the effect of difference in skilled labor endowment to be negative for unskilled-labor intensive industries, that is, $\gamma_2 < 0$, $\gamma_4 > 0$, $\gamma_5 < 0$, and $\gamma_6 < 0$. Our estimates are largely consistent with the literature (see, e.g., Yeaple, 2003a; Alfaro and Charlton, 2009). Consistent with the market access motive, MNCs are found more likely to invest in countries with a larger market size ($\gamma_1 > 0$). Consistent with the comparative advantage motive, MNCs have a greater probability of investing in unskilled labor abundant countries ($\gamma_3 < 0$), especially in unskilled-labor intensive industries ($\gamma_4 > 0$), and trade cost exerts a negative effect on MNCs' investment decisions ($\gamma_2 < 0$ and $\gamma_5 < 0$).[16]

Based on the estimates of equation (5), we obtain and sum, for each host country $\widetilde{c}$ and industry k, the values of $y_{c\widetilde{c}k}$ predicted by market access and comparative advantage factors. To construct predicted MNC activities at a more disaggregated location level, we use the actual share of multinationals in each city to capture cross-city variations in attractiveness (e.g., port access and favorable industrial policies). Multiplying the actual share by $\widehat{y}_{\widetilde{c}k}$ gives $\widehat{y}_{sk}$ for each city s and industry k.

In the alternative specification, we directly estimate MNC activity at a disaggregated region level. To proceed, we re-consider equation (5) to examine MNC activity at the region, instead of country, level and include a series of regional characteristics as additional regressors to capture the effect of regional location fundamentals.

[13] As noted by Helpman (2006), firms' sorting patterns and organization choices are dependent on the characteristics of the firms and the contractual environment (see, e.g., Antras, 2003; Grossman and Helpman, 2002). Existing empirical evidence also suggests that institutional development (such as the rule of law and intellectual property rights) exerts a positive effect on the receipt of foreign investment (Alfaro, Kalemli-Ozcan, and Volosovych, 2008).

[14] Note that the effect of agglomeration forces such as the size of upstream and downstream industries is controlled for in equation (5) by country-industry dummies. Ideally we would like to estimate equation (5) at more disaggregated geographic levels such as cities and provinces, but the explanatory variables in equation (5) are mostly available only at the country level.

[15] See discussion in Santos Silva and Tenreyro (2006) and Head and Ries (2008). We also considered a two-step Heckman selection procedure following Helpman et al. (2008) in which we estimated, respectively, the decision to trade and volume of trade, the results were similar.

[16] Results are suppressed because of space considerations and available upon request.

The main advantage of this specification is that it enables us to examine the role of regional characteristics, such as market size and natural and comparative advantages, in MNCs' location decisions, instead of relying on the role of country characteristics alone and then using a region's share of MNCs as a proxy for regional attractiveness. However, the disadvantage of this specification is the difficulty in obtaining disaggregated regional data for a wide sample of countries.

In the end, we compiled a detailed database of regional characteristics from a number of national sources. For most countries, we were limited by information available to primarily state or province level data. Specifically, for Europe, data was compiled from the Eurostat Regional Database at the NUTS 2 level disaggregation, both to compare with other data and for availability reasons. For other countries, such as US, Australia, Brazil, Canada, China, Japan, Mexico, and South Korea, we used state or province level data. Because of data availability constraints, the regional characteristics systematically available across countries and included in the final sample are income, schooling (educational attainment, percentage of labor with tertiary education), infrastructure (roadways, ports, and airports), and taxes, measured in 2004 or the closest year available (to mitigate potential causality concerns).[17]

Based on this database, we estimate the following equation:

$$\begin{aligned} y_{c\widetilde{c}sk} = {} & \gamma_0 + \gamma_1 marketsize\_size_{c\widetilde{c}} + \gamma_2 distance_{c\widetilde{c}} + \gamma_3 skill\_diff_{c\widetilde{c}s} \\ & + \gamma_4 skill\_diff_{c\widetilde{c}s} \times skillintensity_k + \gamma_5 tariff_{c\widetilde{c}k} + \gamma_6 tariff_{\widetilde{c}ck} \\ & + \gamma_7 tax_{\widetilde{c}s} + \gamma_8 roadway_{\widetilde{c}s} + \gamma_9 port_{\widetilde{c}s} + \gamma_{10} airport_{\widetilde{c}s} + \mu_{ck} + \mu'_{\widetilde{c}k} + \varepsilon_{c\widetilde{c}k}. \end{aligned} \tag{6}$$

where $y_{c\widetilde{c}rs}$ now denotes either the number or the total employment of subsidiaries in country $\widetilde{c}$'s region s and industry k owned by MNCs in country c, $skill\_diff_{c\widetilde{c}s}$ represents the difference in skill endowment, measured by percentage of labor with tertiary education, between the home country and the host region (i.e., $skill_{\widetilde{c}s} - skill_c$), $tax_{\widetilde{c}s}$ is the region's corporate tax level, $roadway$ is the length of roadway in each region s, and $port_{\widetilde{c}s}$ and $airport_{\widetilde{c}s}$ are indicators of ports and airports in the region. Again, we estimate the equation using Poisson quasi-MLE (QMLE) and find estimated parameters to be largely similar to the results from the first specification. In addition, we find regional skill level and infrastructure characteristics to matter significantly in multinationals' location decisions. Similar to the first specification, we then obtain and sum, for

[17] The U.S. data was collected at the state level. Population and education attainment data were collected from the U.S. Census; GDP and income/compensation statistics were collected from the Bureau of Economic Analysis; roadway statistics were from the Federal Highway Administration; employment data was collected from the Bureau of Labor Statistics. Australia data was compiled from the Australian Bureau of Statistics at the state level (ABS). Canada data was collected from Statistics Canada at the provincial level. China data are from the CEIC Data at the provincial level. Brazil data is from IBGE at the state level; data on Brazilian energy production and consumption was from Ministério de Minas e Energia. Mexico data was collected from INEGI at the state level. South Korea data was from KOSIS, collected at the provincial level. Japan data were collected from the Statistics Bureau of Japan at the prefecture level. The remaining data is at the national level, collected from the World Bank. For all regions, port data was from World Port Source, and tax rates were compiled from EY, Deloitte, KPMG, and the World Bank's Doing Business report.

each host country c, region s, and industry k, values of $\widehat{y}_{c\widetilde{c}sk}$ predicted by the market access, comparative advantage, and infrastructure variables.

Step 2 In the second stage, we repeat step 1 of DO's procedure to obtain a geographic distribution function for each pair of industries k and $\widetilde{k}$. We use the predicted levels of MNC activity (either predicted number or total employment of MNCs) in each region and industry (i.e., $\widehat{y}_{sk}$ and $\widehat{y}_{\widetilde{s}\widetilde{k}}$) as the weight when estimating the kernel function. This generates, for each pair of industries, an expected geographic density function based exclusively on the estimated effects of location characteristics including market size, comparative advantage, and trade costs. We compare in Section 6 the role of these characteristics relative to that of agglomeration forces in determining the spatial patterns of multinational firms.

4.2 Agglomeration Economies

In addition to the location fundamentals of MP, agglomeration economies can also affect multinationals' location choices. The advantage of proximity can differ dramatically between multinational and domestic firms and between MNC foreign subsidiaries and MNC headquarters. For instance, multinationals often incur substantial trade costs in sourcing intermediate inputs and reaching downstream buyers. They also face significant market entry costs when relocating to a foreign country because of, for example, limited supplies of capital goods. Further, given their technology intensity, technology diffusion from closely linked industries can be particularly attractive to MNCs.

Agglomeration forces evaluated include (i) vertical production linkages, (ii) externality in labor markets, (iii) externality in capital-good markets, and (iv) technology diffusion. We describe below how each of the above factors is measured in the empirical analysis.

Vertical production linkages Marshall (1890) argued that transportation costs induce plants to locate close to inputs and customers and determine the optimal trading distance between suppliers and buyers. This can be especially true for MNCs given their large volumes of sales and intermediate inputs.[18] Compared to domestic firms, multinationals are often the leading corporations in each industry. Because they tend to be the largest customers of upstream industries as well as the largest suppliers of downstream industries, the input-output relationship between MNCs (e.g., Dell and Intel, Ford and Delphi) can be far stronger than that between average domestic firms.[19]

To determine the importance of customer and supplier relationships in multinationals' agglomeration decisions, we construct a variable, $IOlinkage_{k\widetilde{k}}$, to measure the extent of the input-

[18] For FDI theoretical literature in this area, see, e.g., Krugman (1991), Venables (1996), and Markusen and Venables (2000).

[19] Head, Ries, and Swenson (1995) note, for example, that the dependence of Japanese manufacturers on the "just-in-time" inventory system exerts a particularly strong incentive for vertically linked Japanese firms to agglomerate abroad.

output relationship between each pair of industries. We use the 2002 Benchmark Input-Output (I-O) Data (specifically, the Detailed-Level Make, Use and Direct Requirement Tables) published by the Bureau of Economic Analysis (BEA), and define $IOlinkage_{k\widetilde{k}}$ as the share of industry k's inputs that come directly from industry $\widetilde{k}$, and vice versa. These shares are calculated relative to all input-output flows including those to non-manufacturing industries and final consumers. As supplier flows are not symmetrical, we take either the maximum or the mean of the input and output relationships for each pair of industries.

Externality in labor markets Agglomeration can also yield benefits through external scale economies in labor markets. Because firms' proximity to one another shields workers from the vicissitudes of firm-specific shocks, workers in locations in which other firms stand ready to hire them are often willing to accept lower wages.[20] Externalities can also occur as workers move from one job to another. This is especially true between MNCs which are characterized by similar skill requirements and large expenditures on worker training. MNCs can have a particularly strong incentive to lure workers from one another because the workers tend to receive certain types of training that are well suited for working in most multinational firms (business practices, business culture, etc.).[21]

To examine labor market pooling forces, we follow EGK in measuring each industry pair's similarity in occupational labor requirements. We use the Bureau of Labor Statistics' (BLS) 2006 National Industry-Occupation Employment Matrix (NIOEM), which reports industry-level employment across detailed occupations (e.g., Assemblers and Fabricators, Metal Workers and Plastic Workers, Textile, Apparel, and Furnishings Workers, Business Operations Specialists, Financial Specialists, Computer Support Specialists, and Electrical and Electronics Engineers). We convert occupational employment counts into occupational percentages for each industry, map the BLS industries to the SIC3 framework, and measure each industry pair's labor similarity, $labor_{k\widetilde{k}}$, using the correlation in occupational percentages.

Externality in capital-good markets External scale economies can also arise in capital-good markets. This force has particular relevance to multinational firms given their large involvement in capital-intensive activities as show in Alfaro and Chen (2014).[22] Geographically concentrated industries offer better support to providers of capital goods (e.g. producers of specialized components and providers of machinery maintenance) and reduce the risk of investment (due to, for example, the existence of resale markets).[23] Local expansion of capital intensive

[20]This argument was formally considered in Marshall (1890), Krugman (1991), and Helsley and Strange (1990).

[21]The flow of workers can also lead to technology diffusion, another Marshallian force discussed below.

[22]See Alfaro and Hammel (2007) for evidence on capital flows and capital goods imports.

[23]Agglomeration can also induce costs by, for example, increasing labor and land prices. Like benefits, these costs can be potentially greater for industries with similar labor and capital-good demand, in which case the estimated parameters of the variables would represent the net effect of similar factor demand structures on agglomeration decisions.

activities can consequently lead to expansion of the supply of capital goods, thereby exerting a downward pressure on costs.

To evaluate the role of capital-good market externalities, we construct a new measure of industries' similarity in capital-good demand using capital flow data from BEA. The capital flow table (CFT), a supplement to the 1997 Benchmark Input-Output (I-O) accounts, shows detailed purchases of capital goods (e.g., motors and generators, textile machinery, mining machinery and equipment, wood containers and pallets, computer storage devices, and wireless communications equipment) by industry. We measure each using industry pairs' similarity in capital-good demand structure, denoted by $capitalgood_{k\widetilde{k}}$, using the correlation of investment flow vectors.[24] Industry pairs that exhibit the strongest correlation in capital-good demand include SIC 381 (Search, Detection, Navigation, Guidance, Aeronautical, and Nautical Systems) and SIC 387 (Watches, Clocks, Clockwork Operated Devices, and Parts), SIC 202 (Dairy Products) and SIC 206 (Sugar and Confectionery Products), SIC 326 (Pottery and Related Products) and SIC 328 (Cut Stone and Stone Products), and SIC 221 (Broadwoven Fabric Mills, Cotton) and SIC 228 (Yarn and Thread Mills).

Technology diffusion A fourth motive relates to the diffusion of technologies. Technology can diffuse from one firm to another through movement of workers between companies, interaction between those who perform similar jobs, or direct interaction between firms through technology sourcing. This has been noted by Navaretti and Venables (2006), who predict that MNCs may benefit from setting up affiliates in proximity to other MNCs with advanced technology (e.g., "so-called centers of excellence"). The affiliates can benefit from technology spillovers, which can then be transferred to other parts of the company.

To capture this agglomeration force, we construct a proxy of technology diffusion frequently considered in the knowledge spillover literature (see, e.g., Jaffe et al., 2000; EGK), using patent citation flow data taken from the NBER Patent Database. The data, compiled by Hall et al. (2001), include detailed records for all patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to December 1999. Each patent record provides information about the invention (e.g., technology classification, citations of prior art) and inventors submitting the application (e.g., name and city). We construct the technology diffusion variable, that is, $technology_{k\widetilde{k}}$, by measuring the extent to which technologies in industry k cite technologies in industry $\widetilde{k}$, and vice versa.[25] In practice, there is little directional difference in $technology_{k\widetilde{k}}$ due to the extensive number of citations within a single technology field. We obtain both maximum and mean for each set of pairwise industries.

Constructing the proxies of agglomeration economies using the U.S. industry account data is

[24]Note that this measure captures a different dimension of industry-pair relatedness than vertical production linkages. Unlike vertical production linkages, industry-pair correlations in capital-good demand reflect industry pairs' similarity in capital-good demand and, thus, scope for externality in capital-good markets.

[25]The concordance between the USPTO classification scheme and SIC3 industries is adopted in the construction of the variable.

motivated by three considerations. First, compared to firm-level input-output, factor demand, or technological information, which is typically unavailable, industry-level production, factor and technology linkages reflect standardized production technologies and are relatively stable over time, limiting the potential for the measures to endogenously respond to MNC agglomeration.[26] Second, using the U.S. as the reference country while our analysis covers multinational activity around the world further mitigates the possibility of endogenous production, factor, and technology linkage measures, even though the assumption that the U.S. production structure carries over to other countries could potentially bias our empirical analysis against finding a significant relationship. Third, the U.S. industry accounts are more disaggregated than most other countries', enabling us to dissect linkages between disaggregated product categories.

Table A.1 reports the summary statistics of industry-level control variables. Table A.2 presents the correlation matrix. For example, the correlation between industry-pair production linkages and similarity in capital-good demand is about 0.19, the correlation between production linkages and technology diffusion is about 0.29. The table also shows the mean and maximum measures of production linkages and technology diffusion to be highly correlated.[27]

5 Data: The WorldBase Database

Our empirical analysis employs a unique worldwide establishment dataset, WorldBase, that covers more than 43 million public and private establishments in more than 100 countries and territories (see Alfaro and Chen, 2014). WorldBase is compiled by Dun & Bradstreet (D&B), a leading source of commercial credit and marketing information since 1845, from a wide range of sources.[28] All information collected by D&B is verified centrally via a variety of manual and automated checks.[29]

D&B's WorldBase is, in our view, an ideal data source for the research question proposed in this study offering several distinct advantages over alternative data sources. First, its broad cross-country coverage enables us to examine agglomeration on a global and continuous scale. Examining the global patterns of agglomeration allows us to offer a systematic perspective that takes into account nations at various stages of development. Viewing agglomeration on a continuous scale is important in light of the increasing geographic agglomeration occurring across regional and country borders.

Second, the database reports detailed information for multinational and non-multinational,

[26]Concerns surrounding the endogeneity of agglomeration economies are further discussed and analyzed in section 7; see also discussion in Alfaro and Chen (2014).

[27]We used the mean values in our analysis, but obtained similar results when we used the maximum measure.

[28]For more information, see: http://www.dnb.com/us/about/db_database/dnbinfoquality.html. The dataset employed in this chapter was acquired from D&B with disclosure restrictions.

[29]Alfaro and Charlton's (2009) use the data to study vertical and horizontal activities of multinationals, Alfaro and Chen (2012) MNCs reactions to the Global Financial crisis. The data has also been used to analyze vertical integration decisions in Alfaro, Conconi, Fadinger and Newman (2016) and global integration choices in Alfaro, Antràs , Chor, and Conconi (2015)

offshore and headquarters establishments. This makes it possible to compare agglomeration patterns across different types of establishments and investigate how the economic geography of production evolves with forms of firm organization.

Third, the WorldBase database reports the physical address and postal code of each plant, whereas most existing datasets report business registration addresses. Existing studies have tended to use distance between administrative units, such as state distances, as a proxy for distance of establishments. In doing so, establishments proximate in actual distance but separated by administrative boundaries (e.g., San Diego and Phoenix) can be considered dispersed. Conversely, establishments far in distance but located in the same administrative unit (e.g., San Diego and San Francisco) can be counted as agglomeration.

We obtain latitude and longitude codes for each establishment using a geocoding software (GPS Visualizer). This software uses Yahoo's and Google's Geocoding API services, well known as the industry standard for transportation data. It provides more accurate geocode information than most alternative sources. The geocodes are obtained in batches and verified for precision. We apply the Haversine formula to the geocode data to compute the great circle distance between each pair of establishments.[30] The distance and the trade cost information is used to construct an index of agglomeration following the empirical methodology described in Section 4.

MNC Establishment Data Our main empirical analysis is based on MNC manufacturing establishments in 2005. WorldBase reports, for each establishment in the dataset, detailed information on location, ownership, and activities. In this chapter we use industry information including the four-digit SIC code of the primary industry in which each establishment operates; (ownership information including headquarters, domestic parent, global parent, and position in the hierarchy (branch, division, headquarters); detailed location information for both establishment and headquarters; and operational information including sales, employment, and year started.

An establishment is deemed as an MNC foreign subsidiary if it satisfies two criteria: (i) it reports to a global parent firm, and (ii) the headquarters or the global parent firm is located in a different country. The parent is defined as an entity that has legal and financial responsibility for another establishment.[31] We drop establishments with zero or missing employment values and industries with fewer than 10 observations.[32]

Our final sample includes 32,427 MNC offshore manufacturing plants. Top industries include

[30]To account for other forms of trade barriers, such as border, language, and tariffs, we further estimated a more comprehensive measure of trade cost between each pair of plants. Results available upon request.

[31]There are, of course, establishments that belong to the same multinational family. Although separately examining the interaction of these establishments is beyond the focus of this chapter, we expect the Marshallian forces to have a similar effect here. For example, subsidiaries with an input-output linkage should have incentives to locate near one another independent of ownership. See Yeaple (2003b) for theoretical work in this area and Chen (2011) for supportive empirical evidence. One can use a similar methodology (estimating geographic distributions of establishments that belong to the same firm and comparing them with distributions of counterfactuals) to study intra-firm interaction (see Duranton and Overman, 2008).

[32]Requiring positive employment helps to exclude establishments registered exclusively for tax purposes.

Electronic Components and Accessories (367), Miscellaneous Plastics Products (308), Motor Vehicles and Motor Vehicle Equipment (371), General Industrial Machinery and Equipment (356), Laboratory Apparatus and Analytical, Optical, Measuring, and Controlling Instruments (382), Drugs (283), Metalworking Machinery and Equipment (354), Construction, Mining, and Materials Handling (353), and Special Industry Machinery except Metalworking (355). Top host countries include China, the United States, United Kingdom, Canada, France, Poland, the Czech Republic, and Mexico. More than 20 percent of pairs of multinationals located within 200 km are in different countries. The percentage rises to 45 percent at 400 km and 70 percent at 800 km. This is not surprising given countries' growing participation in regional trading blocs and rapid declines in cross-border trade costs.

Domestic Plant Data Conducting an empirical analysis of all domestic manufacturing plants is infeasible given the size of the entire WorldBase dataset and computational intensity of the procedure. Consequently, to keep the analysis feasible, we adopt a random sampling strategy. For each SIC 3-digit industry with more than 1,000 observations, we obtain a random sample of 1,000 plants. For industries with fewer than 1,000 observations, we include all domestic plants. This yields a final sample of 127,897 domestically owned plants.

6 Assessing the Roles of Location Fundamentals and Agglomeration Economies

We now examine the roles of location fundamentals and agglomeration economies in explaining the pairwise-industry agglomeration of MNCs and how the effects might differ across multinational foreign subsidiaries, domestic plants, and multinational headquarters.

Formally, we estimate the following empirical specification:

$$\begin{aligned} agglomeration_{k\widetilde{k}}(T) &= \alpha_K + \beta_1 fundamentals_{k\widetilde{k}} \\ &\quad + \beta_2 IOlinkage_{k\widetilde{k}} + \beta_3 labor_{k\widetilde{k}} + \beta_4 capitalgood_{k\widetilde{k}} + \beta_5 technology_{k\widetilde{k}} + \varepsilon_{ij}, \end{aligned} \tag{7}$$

where $agglomeration_{k\widetilde{k}}(T)$ is the agglomeration index of industry pairs k and $\widetilde{k}$ at threshold distance T (relative to the counterfactuals) and the right-hand side includes (i) the agglomeration patterns predicted by MP location fundamentals ($fundamentals_{k\widetilde{k}}$) based on the two specifications considered in Section 4.1, and (ii) proxies for agglomeration forces described in Section 4.2 consisting of input-output linkages ($IOlinkage_{k\widetilde{k}}$), labor- and capital-good market similarities ($labor_{k\widetilde{k}}$ and $capitalgood_{k\widetilde{k}}$), and technology diffusion ($technology_{k\widetilde{k}}$). In addition to the location fundamentals and the agglomeration economies considered above, multinationals might also agglomerate because of factors like shared natural advantage (e.g., climate) and externality in institutional and physical infrastructure investment. We account for these factors with both the location fundamental measures and an industry fixed effect. Specifically, we include α_K, a vector of industry dummies that takes the value of 1 if either industry k or $\widetilde{k}$ corresponds to a given industry, and zero otherwise. These industry dummies control for all industry-specific factors and agglomeration patterns. Summary statistics for MNC and domestic agglomeration indices are reported in Tables 1.

6.1 MNC Offshore Agglomeration

We consider first the agglomeration of MNC foreign subsidiaries. Table 2 reports the regression results based on measure 1 of location fundamentals. Agglomeration forces including vertical production linkages, capital-good market correlation, and technology diffusion all play a significant role and display the expected signs.[33] For example, at 200 km a 100-percentage-point increase in the level of technology diffusion, that is, the percentage of patent citations between two industries, leads to a 0.6-percentage-point increase in the level of the agglomeration index between industries. This is equivalent to increasing the average (0.2) by a factor of 3. The location fundamental variable is significant at 1600 km, influencing the spatial patterns of MNCs at a relatively aggregate geographic level.

[Table 2 about here]

The lower panel of Table 2 reports the normalized beta coefficients.[34] Comparing the standardized coefficients of agglomeration forces, we find the effects of technology diffusion and capital-good market correlation to outweigh that of vertical production linkages, which suggests that, given the technology and capital intensive characteristics of multinational firms, it is important to take into account not only vertical production linkages but also technology and capital-good market externalities in explaining MNCs' offshore agglomeration. The parameter of labor-market correlation is insignificant in the multivariate regressions.[35]

Comparing the estimates across distance thresholds, we find that at more aggregate geographic levels, the impact of technology diffusion diminishes and the effect of capital-good market externalities rises while the role of vertical production linkages remains mostly constant. The stronger effect of technology diffusion at shorter distance levels suggests that, compared

[33] In univariate regression results for each of our main variables, all the agglomeration variables were found to be highly significant across the different distance threshold levels. The estimated effects also exhibited expected signs. Across agglomeration forces, capital-good market correlation had the greatest impact across all distance thresholds, followed by labor-demand correlation, technology diffusion, and production linkages. Tables showing univariate results are suppressed from the chapter due to space considerations but available upon request.

[34] Standardized coefficients enable us to compare the changes in the outcomes associated with the metric-free changes in each covariate.

[35] Excluding the capital-good market correlation variable, we found the technology diffusion and production linkage variables to remain positive and significant and the labor correlation coefficient to remain insignificant. This result suggests that the capital-good variable is capturing agglomeration incentives not represented by the other variables.

to the other agglomeration economies, benefits from technology diffusion tend to be localized geographically.[36]

Estimation results based on measure 2 of location fundamentals are reported in Table 3. The estimated parameters of agglomeration economies remain largely similar to Table 2. The location fundamental variable, obtained from the regional-level specification, now exerts a significant effect on the agglomeration of multinational foreign subsidiaries at both 400 and 800 km. Comparing the relative importance of location fundamentals and agglomeration economies, we find the effect of location fundamentals to be outweighed by the effect of the cumulative effect of agglomeration forces in Table 3. At 400 km, a one-standard-deviation increase in location fundamentals leads to a 0.025-standard deviation increase in the level of agglomeration, while the cumulative effect of agglomeration forces is 0.076 standard deviation.[37]

[Table 3 about here]

We have examined MNC offshore agglomeration thus far using subsidiary as the unit of observation. We now take into account the different employment sizes of multinational subsidiaries, which essentially treats the worker as the unit of observation and measures the level of agglomeration among workers. This exercise, by differentiating the agglomeration incentives between individual establishments and workers, has implications for policy making targeted at influencing the geographic distribution of workers.

Tables 4 and 5 reports the estimates based on the two measures of location fundamentals. Note that in contrast to Tables 2 and 3, in which labor market correlation does not exert a significant effect, multinational subsidiaries in industries with greater potential labor market externalities exhibit significantly higher level of employment agglomeration. Technology diffusion, another force of agglomeration that involves close labor interaction and mobility, also plays a significant role in explaining the agglomeration of MNC subsidiary workers between industries. In fact, technology spillover appears to be the strongest agglomeration factor at most distance thresholds. Further, at more aggregate geographic levels, the effects of labor market externalities and technology spillovers diminish, while capital-good market correlation exerts a significant and positive effect.

[Tables 4 and 5 about here]

[36]When excluding the location fundamental variable, the coefficients and statistical significance of the agglomeration forces remain largely unchanged.

[37]Comparing Table 2 and Table 3, we also note the normalized parameter of the location fundamental variable to be significantly lower when the variable is constructed based on the regional estimation specification. One possible explanation is that measure 1, constructed based on country-level location characteristics and actual regional share of multinational activity, represents an upper bound of location fundamentals, whereas measure 2, estimated based on observable country and regional characteristics, serves as a lower bound.

6.2 MNC Headquarters Agglomeration

We next examine the determinants of MNC headquarters clusters relative to MNC clusters overseas. To control for the role of location fundamentals in explaining the agglomeration of MNC headquarters, we follow the procedure described in Section 4.1, but obtain the level of MNC activities predicted for each MNC home country, and construct the expected distribution and agglomeration of MNC headquarters following the rest of the procedure.

Table 6 reports the estimation results. All variables except vertical production linkages exert a significant effect. A one-standard-deviation increase in the location fundamental variable is associated with a 0.21 standard-deviation increase in MNC headquarters agglomeration, which suggests an important role for the characteristics of headquarter countries including market size, skilled labor endowment, and access to host countries. At 200 km, both technology diffusion and labor market correlation play a positive and significant role, with a cumulative effect of about 0.06. Beyond 200 km, the effect of labor market becomes insignificant. Again, this result is consistent with the localized feature of labor markets and lower mobility of labor.

[Table 6 about here]

Comparing Table 6 with Table 2, we find that location fundamentals and capital-good market externality exert a stronger effect on MNCs' offshore agglomeration than on the agglomeration of MNC headquarters and, further, input-output relationships affect MNC subsidiaries but not headquarters. These results suggest that MNC subsidiary agglomeration, with their market-seeking and input-sourcing focuses, is more influenced by market access and comparative advantage motives, capital-good market externalities, and vertical production linkages, whereas agglomeration of headquarters, with their specialization in R&D, management, and the provision of other services, is more influenced by technology diffusion than by production linkages.

6.3 Comparing the Agglomeration of MNC Offshore and Domestic Plants

Having established the agglomeration patterns of MNC foreign subsidiaries, we now investigate how the role of agglomeration forces varies systematically between multinational and non-multinational plants. Specifically, we evaluate how the role of location fundamentals and agglomeration economies affects MNCs relative to domestic plants by estimating the following equation:

$$\begin{aligned}&agglomeration^m_{k\widetilde{k}}(T) - agglomeration^d_{k\widetilde{k}}(T)\\&= (\beta^m_1 - \beta^d_1) fundamentals_{k\widetilde{k}} + (\beta^m_2 - \beta^d_2) IOlinkage_{k\widetilde{k}} + (\beta^m_3 - \beta^d_3) labor_{k\widetilde{k}}\\&+ (\beta^m_4 - \beta^d_4) capitalgood_{k\widetilde{k}} + (\beta^m_5 - \beta^d_5) technology_{k\widetilde{k}} + \varepsilon_{ij},\end{aligned} \tag{8}$$

where $agglomeration^m_{k\widetilde{k}}(T) - agglomeration^d_{k\widetilde{k}}(T)$ represents the difference between the MNC and domestic pairwise-industry agglomeration indices, and the coefficient vector $\beta^m - \beta^d$ represents the difference in the effects of the covariates on multinational foreign subsidiaries and domestic plants.

[Tables 7 and 8 about here]

The results based on the two measures of location fundamentals are reported in Tables 7 and 8. We find that proxies for capital-good market externalities and technology diffusion exert a stronger effect on multinationals than on domestic plants in same industry pairs. The role of the input-output relationship is not significantly different between the two at disaggregated geographic levels, but is significantly stronger for multinationals at more aggregate geographic levels (e.g., 800 km). Interestingly, potential externalities in labor market, captured by industry-pair similarity in labor demand, exert a greater effect on the agglomeration of domestic plants than the agglomeration of multinational foreign subsidiaries. Location fundamental variables including market size, comparative advantage, and infrastructure also have a greater role in the agglomeration patterns of domestic plants.

These findings are consistent with the characteristics of multinational firms: relative to their domestic counterparts, multinationals exhibit greater participation in knowledge- and capital-intensive activities and would thereby enjoy stronger agglomeration economies in technology and capital-good markets. Externalities such as technology diffusion and capital-good market scale economies thus provide multinational subsidiaries greater incentives to agglomerate with one another relative to domestic plants. Domestic plants, in contrast, place a greater emphasis on fundamental location characteristics such as market size, production cost, and infrastructure and labor market considerations.

7 Additional Econometric Analysis

A potential concern with our analysis thus far is that the agglomeration economy measures might endogenously reflect the agglomeration patterns of multinational firms. For example, the input-output linkage between the apparel and cotton industries may reflect not just the inherent characteristics of apparel manufacturing, but also the agglomeration of the two industries due, for example, to availability of raw materials leading apparel manufacturers to favor cotton over other types of fabrics. Similarly, the technology spillover between the telecommunication and computer industries might be due not only to the intrinsic technological relationship between the two industries, but also to a historical factor that led the two industries to locate together and subsequently become familiar with each other's technologies.

This concern is mitigated in our work by three factors. First, our analysis controls for the role of location fundamentals and industry-specific characteristics. This enables us to separate industries' geographic concentration due to location attractiveness from agglomeration activities driven by agglomeration economies. Second, our measures of agglomeration economies are constructed using U.S. industry account data while the chapter examines global agglomeration patterns. U.S. industries' input-output linkages, factor market correlations, and technology spillovers are not very likely a result of agglomeration around the world. Third, the focus on MNCs reduces the possibility of reverse causation, as MNCs constitute a small subset of firms in each industry and the agglomeration economy measures are built with industry wide data that include information on domestic firms.[38]

We nevertheless perform an additional exercise to further alleviate concerns about endogeneity. Because the global agglomeration patterns of multinational firms include the agglomeration of MNCs in the United States, we examine regional agglomeration for which the U.S. is excluded. If U.S. domestic industry-pair relationships are affected by the agglomeration of MNCs in the United States, then one would expect the former to be less likely affected by the agglomeration of MNCs located in other regions such as Europe.[39] In this case, the agglomeration economy measures constructed with U.S. industry account data are orthogonal to the agglomeration patterns observed in Europe.[40]

We proceed by repeating the procedure described in Section 4.1 to construct the agglomeration indices for MNCs located in Europe. These indices capture the degree to which MNCs in a given industry pair agglomerate in Europe at various threshold distances.

[Table 9 about here]

The results are reported in Table 9. We find the estimates to be qualitatively similar to those reported in Tables 2 and 3.[41] Multinational subsidiaries in industries with greater labor market correlation and technology spillover are found to have a higher level of agglomeration, especially at the 200 and 400 km levels. Input-output production linkage and capital-good market correlation also exert a significant effect on the agglomeration of MNCs in Europe. Consistent with the earlier results, we find the effects of labor market externalities and technology spillovers to diminish at more aggregate geographic levels. Further, labor market externality appears to be the strongest agglomeration force at disaggregated distance levels.

8 Conclusion

The emergence of new multinational clusters is one of the most notable phenomena in the process of globalization. Multinationals follow distinctively different agglomeration patterns offshore than their domestic counterparts (Alfaro and Chen, 2014). We examine in this chapter the

[38] Alfaro and Chen (2014) further explore the process of agglomeration.

[39] On regional integration and concentration of MNCs, see also Chen (2009).

[40] Using another country's data to instrument the agglomeration economy variables would not alleviate the potential for endogeneity in our analysis because it would face issues similar to the U.S. data. Using the U.S. agglomeration economy measures to predict the agglomeration patterns in a non-U.S. region would, however, mitigate the possibility of reverse causation and help identify the causal effects of agglomeration forces.

[41] Because we are now examining regional, instead of global, agglomeration, we consider only threshold distances up to 800 km.

relative importance of agglomeration forces versus location fundamentals in MNCs' offshore as well as headquarter geographic patterns. Our analysis, using a worldwide plant-level dataset and a novel index of agglomeration, yields a number of new insights into the economic geography of multinational production.

MP location fundamentals, although playing a significant and important role in explaining the agglomeration of multinational firms, are not the only driving force. In addition to market access and comparative advantage motives, multinationals' location choices are significantly affected by agglomeration economies including not only vertical production linkages but also technology diffusion and capital-market externalities. Further, the importance of location fundamentals and agglomeration economies varies significantly between MNCs' offshore agglomeration and the agglomeration of MNC headquarters and domestic plants. For example, MNCs' offshore plants are significantly more influenced than non-MNC plants by capital-good market and technological agglomeration factors.

Our results convey implications central to academic and policy debates on FDI. The agglomeration of economic activity, as long recognized by regional and urban economists and economic historians, is one of the salient features of economic development. An extensive body of research examines the distribution of population and production across space and the economic characteristics and effects of spatial concentrations. Understanding the emerging spatial concentrations of multinational production around the world and the driving forces behind these new concentrations in comparison to those of their domestic counterparts is crucial for designing and improving policies. Growing evidence suggests that multinationals play a significant role in the performance of local economies, raising local wages (see, for example, Aitken, Harrison, and Lipsey, 1996), spillovers (Javorcik, 2004) and reallocation (see, for example Alfaro and Chen, 2015). Evidence has shown positive effects of FDI on host country's growth conditional on local conditions (Alfaro et al., 2004, 2010) and resilience to external shocks (Alfaro and Chen, 2012). Recognizing these effects, many countries, including both FDI source and destination nations, have long offered lucrative incentives to MNCs in the hope of building and sustaining industrial clusters. Understanding the location interdependence of multinational firms and how they agglomerate with one another is critical to designing these economic policies.

.

References

[1] Aitken, Brian, Harrison, Ann and Lipsey, Robert (1996) Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela, and the United States. *Journal of International Economics* 40(3-4): 345-371.

[2] Alfaro, Laura, Chanda, Areendam, Kalemli-Ozcan, Sebnem, Sayek, Selin (2004) FDI and Economic Growth: the Role of Local Financial Markets. *Journal of International Economics* 64 (1), 89-112.

[3] Alfaro, Laura, Areendam, Kalemli-Ozcan, Sebnem, Sayek, Selin. 2010. "Does Foreign Direct Investment Promote Growth? Exploring the Role of Financial Markets on Linkages." *Journal of Development Economics* 91(2), 242-256.

[4] Alfaro, Laura and Charlton, Andrew (2009) Intra-Industry Foreign Direct Investment. *American Economic Review* 99(5): 2096-2119.

[5] Alfaro, Laura and Chen, Maggie X. (2012) Surviving the Global Financial Crisis: Foreign Ownership and Establishment Performance. *American Economic Journal: Economic Policy* 4: 30-55.

[6] Alfaro, Laura and Chen, Maggie X. (2014) The Global Agglomeration of Multinational Firms. *Journal of International Economics* 99(2): 263-276

[7] Alfaro, Laura and Chen, Maggie X. (2015) Selection and Market Reallocation: Productivity Gains from Multinational Production. Harvard Business School Working Paper No. 12-111.

[8] Alfaro, Laura, Antràs, Pol, Chor, Davin and Conconi, Paola(2015) Internalizing Global Value Chains: A Firm-Level Analysis. NBER Working Paper 21582.

[9] Alfaro, Laura, Conconi, Paola, Fadinger, Harald, Newman, Andrew (2016) Do Prices Determine Vertical Integration? *Review of Economics Strudies* 83(3): 855-888.

[10] Alfaro, Laura and Hammel, Eliza (2007) Capital Flows and Capital Goods *Journal of International Economics* 72(1): 128-150.

[11] Alfaro, Laura, Kalemli-Ozcan, Sebnem, and Volosovych, Vadym (2008) Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigation. *Review of Economics and Statistics* 90: 347-368.

[12] Antràs, Pol (2003) Firms, Contracts and Trade Structure. *Quarterly Journal of Economics* 118: 1375-1418.

[13] Antràs, Pol and Helpman, Elhanan (2004) Global Sourcing. *Journal of Political Economy* 112(3): 552-580.

[14] Antràs, Pol and Helpman, Elhanan (2008) ”Contractual Frictions and Global Sourcing,” in Elhanan Helpman, Dalia Marin, and Thierry Verdier (Eds.), *The Organization of Firms in a Global Economy*, Cambridge: Harvard University Press, forthcoming.

[15] Bobonis, Gustavo J. and Shatz, Howard J. (2007) Agglomeration, Adjustment, and State Policies in the Location of Foreign Direct Investment in the United States. *Review of Economics and Statistics* 89(1): 30-43.

[16] Carr L. David, Markusen, James R. and Maskus, Keith E. (2001) Estimating the Knowledge-Capital Model of the Multinational Enterprise. *American Economic Review* 91(3): 693-708.

[17] Chen, Maggie X. (2009) Regional Economic Integration and Geographic Concentration of Multinational Firms. *European Economic Review* 53(3): 355-375.

[18] Chen, Maggie X. (2011) Interdependence in Multinational Production Networks. *Canadian Journal of Economics* 44(3): 930-956.

[19] Chen, Maggie X. and Moore, Michael (2010) Location Decision of Heterogeneous Multinational Firms. *Journal of International Economics* 80(2): 188-199.

[20] Duranton, Gilles and Kerr, William (2015) The Logic of Agglomeration. NBER Working Papers 21452.

[21] Duranton, Gilles and Overman, Henry (2005) Testing for Localization Using MicroGeographic Data. *Review of Economic Studies* 72(4): 1077-1106.

[22] Duranton, Gilles and Overman, Henry (2008) Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data. *Journal of Regional Science* 48(1): 213-243.

[23] Duranton, Gilles and Puga, Diego (2004) Micro-foundations of urban agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, Vol. 4. Amsterdam: North-Holland, 2063-2117.

[24] Ellison, Glenn and Glaeser, Edward (1997) Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5): 889-927.

[25] Ellison, Glenn, Glaeser, Edward and Kerr, William (2010) What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100: 1195–1213.

[26] Grossman, Gene M. and Helpman, Elhanan (2002) Integration Versus Outsourcing In Industry Equilibrium. *Quarterly Journal of Economics* 117(1): 85-120.

[27] Hall, Bronwyn, Jaffe, Adam and Trajtenberg, Manuel (2001) The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Paper 8498.

[28] Head, Keith and Mayer, Theirry (2004) The Empirics of Agglomeration and Trade. In: Henderson, J.V., and Thisse, J.-F. (Eds.), *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: Elvieser, pp. 2609-2669.

[29] Head, Keith, Ries, John and Swenson, Deborah (1995) Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States. *Journal of International Economics* 38(3-4): 223-247.

[30] Head, Keith and Ries, John (2008) FDI as an Outcome of the Market for Corporate Control: Theory and Evidence. *Journal of International Economics* 74(1): 2-20.

[31] Helpman, Elhanan (1984) A Simple Theory of Trade with Multinational Corporations. *Journal of Political Economy* 92(3): 451-471.

[32] Helpman, Elhanan (2006) Trade, FDI and the Organization of Firms. *Journal of Economic Literature* 44(3): 589-630.

[33] Helpman, Elhanan, Melitz, Marc and Rubinstein, Yona (2008) Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics* 123(2): 441-487.

[34] Helpman, Elhanan, Melitz, Marc, and Yeaple, Stephen (2004) Export versus FDI with Heterogeneous Firms. *American Economic Review* 94 (1): 300-316.

[35] Helsley, Robert W. and Strange, William C. (1990) Matching and Agglomeration Economies in a System of Cities. *Regional Science and Urban Economics* 20(2): 189-212.

[36] Jaffe, Adam, Trajtenberg, Manuel and Fogarty, Michael (2000) Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economic Review Paper and Proceedings* 90: 215-218.

[37] Javorcik, Beata S. (2004) Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. *American Economic Review* 94(3): 605-627.

[38] Krugman, Paul (1991) Increasing Returns and Economic Geography. *Journal of Political Economy* 99(3): 483-499.

[39] Markusen, James (1984) Multinationals, Multi-Plant Economies, and the Gains from Trade. *Journal of International Economics* 16(3-4): 205-226.

[40] Markusen, James and Venables, Anthony J. (1998) Multinational Firms and the New Trade Theory. *Journal of International Economics* 46(2): 183-203.

[41] Markusen, James and Venables, Anthony J. (2000) The Theory of Endowment, Intra-industry and Multi-national trade. *Journal of International Economics* 52(2): 209-234.

[42] Markusen, James (2002) Multinational Firms and the Theory of International Trade. Cambridge: MIT Press.

[43] Marshall, Alfred (1890) *Principles of Economics*. London: MacMillan and Co.

[44] Navaretti, Giorgio Barba and Venables, Anthony J. (2006) *Multinational Firms in the World Economy*. Princeton University Press.

[45] Nunn, Nathan (2007) Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade. *The Quarterly Journal of Economics* 122(2): 569-600.

[46] Ottaviano, Gianmarco I. P. and Puga, Diego, (1998) Agglomeration in the Global Economy: A Survey of the "New Economic Geography." *World Economy* 21(6): 707-731.

[47] Ottaviano, Gianmarco and Thisse, Jacques-François (2004). Agglomeration and economic geography. In: Henderson, J. Vernon, and Thisse, Jacques-François. (Eds.), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, pp.2563-2608.

[48] Overman, Henry and Puga, Diego (2009) Labour Pooling As a Source of Agglomeration: An Empirical Investigation. CEPR Discussion Papers 7174.

[49] Puga, Diego (2010) The Magnitude and Causes of Agglomeration Economies. *Journal of Regional Science* 50(1): 203-219.

[50] Redding, Stephen (2010) The Empirics of New Economic Geography. *Journal of Regional Science* 50(1): 297-311.

[51] Redding, Stephen (2011) Economic Geography: a Review of the Theoretical and Empirical Literature. *The Palgrave Handbook of International Trade*, Chapter 16.

[52] Rosenthal, Stuart and Strange, William (2001) The Determinants of Agglomeration. *Journal of Urban Economics* 50(2):191–229.

[53] Rosenthal, Stuart and Strange, William (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson, J. Vernon, Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier, Amsterdam, pp. 2119-2171.

[54] Santos Silva, J. M. C. and Tenreyro, Silvana (2006) The Log of Gravity. *Review of Economics and Statistics* 88(4): 641-658.

[55] United Nations Conference on Trade and Development (2015) *Trade and Investment Report*, New York: United Nations.

[56] Venables, Anthony (1996) Equilibrium Locations of Vertically Linked Industries. *International Economic Review* 37(2): 341-359.

[57] Yeaple, Stephen (2003a) The Role of Skill Endowments in the Structure of U.S. Outward Foreign Direct Investment. *Review of Economics and Statistics* 85(3): 726-734.

[58] Yeaple, Stephen (2003b) The Complex Integration Strategies of Multinational Firms and Cross-Country Dependencies in the Structure of Foreign Direct Investment. *Journal of International Economics* 60(2): 293-314.

[59] Yeaple, Stephen (2009) Firm Heterogeneity and the Structure of U.S. Multinational Activity: An Empirical Analysis. *Journal of International Economics* 78(2): 206-215.

Table 1: Descriptive Statistics for MNC and Domestic Agglomeration Indices

	Obs.	Mean	Std. Dev.	Min.	Max.
Subsidiaries (Percentage Points)					
Threshold (T) = 200 km	8004	0.099	0.239	0.000	3.060
T = 400 km	8004	0.219	0.522	0.000	6.631
T= 800 km	8004	0.520	1.206	0.000	14.419
T= 1600 km	8004	1.028	2.357	0.000	23.941
Domestic Plants (Percentage Points)					
Threshold (T) = 200 km	8004	0.102	0.289	0.000	4.012
T = 400 km	8004	0.235	0.545	0.000	7.935
T= 800 km	8004	0.550	1.384	0.000	16.539
T= 1600 km	8004	1.210	2.424	0.000	26.340
Subsidiaries Workers (Percentage Points)					
Threshold (T) = 200 km	8004	0.095	0.274	0.000	2.997
T = 400 km	8004	0.194	0.528	0.000	5.553
T= 800 km	8004	0.418	1.038	0.000	10.139
T= 1600 km	8004	0.742	1.853	0.000	17.211
Headquarters (Percentage Points)					
Threshold (T) = 200 km	8004	0.140	0.348	0.000	8.400
T = 400 km	8004	0.325	0.779	0.000	18.198
T= 800 km	8004	0.782	1.772	0.000	39.871
T= 1600 km	8004	1.402	2.987	0.000	44.693

Notes: The agglomeration indices are constructed by comparing the estimated distance kernel function of each industry pair with the 95 percent global confidence band of counterfactual kernel estimators at 200 km, 400 km, 800 km, and 1600 km. All industry pairs (SIC3) are included. See text for detailed descriptions of the variables.

Table 2: Location Fundamentals, Agglomeration Economies, and MNC Subsidiary Agglomeration I

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	0.265*	0.573*	1.331**	2.596**
	(0.147)	(0.306)	(0.656)	(1.296)
Capital Good	0.038***	0.093***	0.241***	0.506***
	(0.014)	(0.032)	(0.066)	(0.139)
Labor	-0.002	-0.015	-0.079	-0.231
	(0.016)	(0.035)	(0.068)	(0.160)
Technology	0.609**	1.178**	2.521**	4.395**
	(0.293)	(0.546)	(1.117)	(2.371)
Location Fundamentals	0.018	0.019	0.020	0.021*
	(0.025)	(0.019)	(0.022)	(0.012)
Obs.	7875	7875	7875	7875
R^2	0.571	0.600	0.627	0.631
	Beta Coefficients			
IO Linkages	0.014	0.014	0.014	0.013
Capital Good	0.035	0.039	0.043	0.046
Labor	-0.002	-0.007	-0.015	-0.023
Technology	0.031	0.027	0.025	0.022
Location Fundamentals	0.266	0.264	0.279	0.333

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 3: Location Fundamentals, Agglomeration Economies, and MNC Subsidiary Agglomeration II

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	0.249**	0.541*	1.252***	2.413***
	(0.112)	(0.302)	(0.222)	(0.576)
Capital Good	0.037**	0.092***	0.237***	0.499***
	(0.017)	(0.017)	(0.092)	(0.153)
Labor	0.001	-0.001	-0.045	0.153
	(0.014)	(0.015)	(0.165)	(0.135)
Technology	0.573***	1.101***	2.330***	3.943*
	(0.161)	(0.458)	(0.343)	(2.560)
Location Fundamentals (Regional)	0.006	0.004***	0.002*	0.001
	(0.007)	(0.001)	(0.001)	(0.003)
Obs.	7875	7875	7875	7875
R^2	0.570	0.600	0.626	0.630
		Beta Coefficients		
IO Linkages	0.013	0.013	0.013	0.012
Capital Good	0.034	0.038	0.042	0.045
Labor	0.004	-0.001	-0.009	-0.015
Technology	0.029	0.025	0.023	0.019
Location Fundamentals (Regional)	0.038	0.025	0.013	0.006

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 4: Location Fundamentals, Agglomeration Economies, and MNC Subsidiary Worker Agglomeration I

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	-0.145	-0.256	-0.272	-0.750
	(0.209)	(0.403)	(0.683)	(1.160)
Capital Good	0.041*	0.109**	0.315***	0.557***
	(0.023)	(0.044)	(0.089)	(0.144)
Labor	0.048*	0.088*	0.120	0.128
	(0.026)	(0.048)	(0.104)	(0.162)
Technology	2.262***	3.957***	6.243***	9.333***
	(0.516)	(0.867)	(1.613)	(2.356)
Location Fundamentals	0.0004***	0.0004***	0.0004***	0.0004**
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Obs.	7875	7875	7875	7875
R^2	0.327	0.327	0.363	0.402
	Beta Coefficients			
IO Linkages	-0.007	-0.006	-0.003	-0.005
Capital Good	0.033	0.045	0.066	0.065
Labor	0.042	0.039	0.027	0.016
Technology	0.100	0.091	0.073	0.061
Location Fundamentals	0.315	0.349	0.390	0.435

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 5: Location Fundamentals, Agglomeration Economies, and MNC Subsidiary Worker Agglomeration II

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	-0.151	-0.269	-0.299	-0.801
	(0.120)	(0.212)	(0.482)	(0.835)
Capital Good	0.040***	0.106***	0.308***	0.544***
	(0.014)	(0.018)	(0.087)	(0.176)
Labor	0.057***	0.107**	0.162*	0.212
	(0.022)	(0.049)	(0.077)	(0.050)
Technology	2.228***	3.885***	6.083***	9.013***
	(0.508)	(0.326)	(1.390)	(2.815)
Location Fundamentals (Regional)	0.002	0.004**	0.007*	0.009*
	(0.010)	(0.002)	(0.004)	(0.001)
Obs.	7875	7875	7875	7875
R^2	0.326	0.326	0.363	0.402
		Beta Coefficients		
IO Linkages	-0.007	-0.006	-0.003	-0.005
Capital Good	0.032	0.044	0.064	0.065
Labor	0.049	0.047	0.036	0.026
Technology	0.100	0.089	0.071	0.058
Location Fundamentals (Regional)	0.011	0.027	0.054	0.086

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 6: Location Fundamentals, Agglomeration Economies, and MNC Headquarters Agglomeration

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	0.090	0.156	0.127	0.457
	(0.174)	(0.406)	(0.815)	(1.254)
Capital Good	0.026	0.084**	0.261***	0.459***
	(0.019)	(0.040)	(0.088)	(0.164)
Labor	0.043**	0.064	0.019	-0.085
	(0.021)	(0.044)	(0.104)	(0.180)
Technology	0.793***	1.727***	3.870***	6.935***
	(0.241)	(0.477)	(1.153)	(1.735)
Location Fundamentals	0.022**	0.023***	0.024*	0.019
	(0.009)	(0.009)	(0.013)	(0.018)
Obs.	7875	7875	7875	7875
R^2	0.639	0.65	0.664	0.667
	Beta Coefficients			
IO Linkages	0.003	0.003	0.001	0.002
Capital Good	0.017	0.024	0.032	0.033
Labor	0.030	0.020	0.003	-0.007
Technology	0.028	0.027	0.027	0.028
Location Fundamentals	0.212	0.212	0.208	0.213

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 7: Comparing MNC Subsidiaries with Domestic Plants I

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	0.041	1.081	5.447**	10.876**
	(0.599)	(1.306)	(2.760)	(4.437)
Capital Good	0.162***	0.494***	1.335***	2.383***
	(0.051)	(0.113)	(0.220)	(0.366)
Labor	-0.110**	-0.443***	-1.430***	-2.130***
	(0.049)	(0.112)	(0.231)	(0.410)
Technology	-1.214	2.823*	24.272***	62.572***
	(0.839)	(1.706)	(3.409)	(6.220)
Location Fundamentals	-0.047***	-0.047***	-0.044***	-0.035***
	(0.003)	(0.002)	(0.002)	(0.002)
Obs.	7875	7875	7875	7875
R^2	0.049	0.053	0.064	0.073
	Beta Coefficients			
IO Linkages	0.001	0.008	0.020	0.023
Capital Good	0.047	0.067	0.085	0.086
Labor	-0.034	-0.065	-0.099	-0.084
Technology	-0.020	0.021	0.086	0.126
Location Fundamentals	-0.213	-0.217	-0.219	-0.228

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 8: Comparing MNC Subsidiaries with Domestic Plants II

	T= 200 km	T= 400 km	T= 800 km	T= 1600 km
IO Linkages	-0.023	0.916	5.014**	10.094**
	(0.603)	(1.285)	(2.515)	(4.406)
Capital Good	0.183***	0.536***	1.421***	2.533***
	(0.048)	(0.118)	(0.217)	(0.375)
Labor	-0.264***	-0.774***	-2.136***	-3.419***
	(0.045)	(0.102)	(0.225)	(0.406)
Technology	0.943	1.252**	20.632***	55.824***
	(0.880)	(0.602)	(3.346)	(6.314)
Location Fundamentals (Regional)	-0.011***	-0.010***	-0.025***	-0.454***
	(0.001)	(0.001)	(0.003)	(0.005)
Obs.	7875	7875	7875	7875
R^2	0.012	0.016	0.028	0.034
		Beta Coefficients		
IO Linkages	-0.0004	0.007	0.018	0.021
Capital Good	0.053	0.072	0.090	0.091
Labor	-0.083	-0.114	-0.148	-0.134
Technology	0.007	0.009	0.073	0.112
Location Fundamentals (Regional)	-0.079	-0.089	-0.101	-0.103

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table 9: The Endogeneity of Agglomeration Economy Measures: The Agglomeration Patterns of MNCs in Europe

	T= 200 kms	T= 400 kms	T= 800 kms
IO Linkages	0.104	0.248*	0.454**
	(0.079)	(0.157)	(0.209)
Capital Good	0.008	0.031*	0.044*
	(0.010)	(0.019)	(0.026)
Labor	0.031***	0.032*	0.036
	(0.008)	(0.018)	(0.030)
Technology	0.335**	0.514**	0.715**
	(0.151)	(0.262)	(0.393)
Location Fundamentals	-0.001	-0.004	-0.003
	(0.003)	(0.005)	(0.004)
Obs.	7166	7166	7166
R^2	0.635	0.717	0.853
		Beta Coefficients	
IO Linkages	0.009	0.009	0.008
Capital Good	0.014	0.021	0.014
Labor	0.055	0.023	0.012
Technology	0.030	0.019	0.013
Location Fundamentals	-0.158	-0.087	-0.076

Notes: Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All regressions include industry fixed effect. Normalized beta coefficients in lower panel. See text for detailed descriptions of the variables.

Table A.1: Descriptive Statistics for Agglomeration Economies

	# Obs.	Mean	Std. Dev.	Min.	Max.
Input-Output (IO) Linkages	7875	0.003	0.012	0.000	0.193
Capital Good	7875	0.476	0.209	-0.004	1.000
Labor	7875	0.333	0.227	0.014	1.000
Technology	7875	0.007	0.012	0.000	0.179

Notes: Input-Output (IO) Linkages, Capital Good, Labor, and Technology correspond to the industry-level variables employed to proxy for the various agglomeration economies: vertical production linkages, externalities in factor markets including labor and capital goods, and technology diffusion. Same industry pairs (SIC3) are excluded. See text for detailed descriptions of the variables.

Table A.2: Correlation of Agglomeration Economies

	IO Linkages	IO Linkages (max.)	Capital Good	Labor	Technology	Technology (max.)
IO Linkages	1.000					
IO Linkages (max.)	0.973	1.000				
Capital Good	0.191	0.189	1.000			
Labor	0.232	0.225	0.567	1.000		
Technology	0.291	0.284	0.230	0.331	1.000	
Technology (max.)	0.264	0.257	0.188	0.297	0.976	1.000

Notes: Obs=7875. Both average and maximum measures are obtained for IO linkages and technology diffusion. See text for detailed descriptions of the variables.