Journal of International Economics 94 (2014) 263–276

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

The global agglomeration of multinational firms☆

Laura Alfaro [a,b,1], Maggie Xiaoyang Chen [c,2]

[a] Harvard Business School, Morgan 293, Boston, MA 02163, United States
[b] NBER, United States
[c] Department of Economics, George Washington University, 2115 G ST, NW, #367, Washington, DC 20052, United States

ARTICLE INFO

Article history:
Received 21 January 2013
Received in revised form 3 September 2014
Accepted 29 September 2014
Available online 12 October 2014

JEL classification:
F2
D2
R1

Keywords:
Multinational firm
Agglomeration
Agglomeration economies

ABSTRACT

The explosion of multinational activities in recent decades is rapidly transforming the global landscape of industrial production. But are the emerging clusters of multinational production the rule or the exception? What drives the offshore agglomeration of multinational firms in comparison to the agglomeration of domestic firms? 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 agglomeration and analyze the different patterns underlying the global economic geography of multinational and non-multinational firms. We present new stylized facts that suggest that the offshore clusters of multinationals are not a simple reflection of domestic industrial clusters. 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 we explore the process of agglomeration.

1. Introduction

The explosion of multinational activities in recent decades is rapidly transforming the global landscape of industrial production. But are the emerging clusters of multinational corporations (MNCs) the rule or the exception? What drives the offshore agglomeration of MNCs in comparison to the agglomeration of domestic firms? In this paper, we examine the patterns of the global agglomeration of multinational production—both offshore and at headquarters—in comparison to the agglomeration of domestic firms.

We quantify and characterize the global agglomeration of multinational and domestic firms to establish new insights into how firms of different organizational forms might agglomerate differently. We use the term agglomeration broadly to explore the geographic concentration of production activities.[3] As highlighted in a growing literature led by Helpman et al. (2004) and Antràs and Helpman (2004, 2008), the economic attributes and organizations of multinationals are, by selection, different from those of domestic firms. The greater revenue and productivity, the vertically integrated production, and the higher knowledge- and capital-intensities all suggest that the agglomeration motives of MNC offshore subsidiaries are likely to be different from those of domestic firms.

We use WorldBase, a worldwide plant-level dataset that provides detailed location, ownership, and activity information for over 43 million plants—including multinational and domestic, offshore and headquarters establishments—in more than 100 countries. This dataset makes it possible to compare the agglomeration of different types of establishment. We use the plant-level physical location information to obtain latitude and longitude codes for each establishment and compute the distance between each pair of establishments.

To quantify the agglomeration patterns, we construct an index of agglomeration at both the pairwise industry level and the plant level by extending an empirical methodology introduced by Duranton and Overman (2005) (henceforth, "DO"). The index measures the extent of geographic localization and the spatial scale at which it takes place. It

☆ We are very grateful to the Editor Stephen Yeaple and two anonymous referees for their insightful comments and suggestions. We thank Jim Anderson, Bruce Blonigen, James Harrigan, Keith Head, Tarun Khanna, Jim Markusen, Keith Maskus, Mike Moore, Henry Overman, John Ries, Roberto Samaneigo, Tony Yezer, and seminar and conference participants at Harvard Business School, University of Virginia Darden School of Business, George Washington University, the AEA meeting, the ETSG meeting, the EIIT meeting, and the LACEA Trade, Integration and Growth Meeting for helpful suggestions, William Kerr for discussions on the agglomeration indices and for kindly providing us the patent concordance data, Bill Simpson for advice on computing the indices, John Elder, Elizabeth Meyer, Francisco Pino, and Hillary White for research assistance, Dennis Jacques for help with the D&B dataset, and HBS and GW for financial support.

E-mail addresses: lalfaro@hbs.edu (L. Alfaro), xchen@gwu.edu (M.X. Chen).
[1] Tel.: +1 617 495 7981.
[2] Tel.: +1 202 994 0192.

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

http://dx.doi.org/10.1016/j.jinteco.2014.09.001
0022-1996/

first estimates the actual density function of distance between MNC establishments and then compares that density function with the counterfactual. In our main analysis, we use the distance density function of domestic establishments in the same industry as the counterfactual to control for the role of location fundamentals that affect both MNC and domestic plants. The index thus quantifies the extent to which MNC establishments are more or less likely to agglomerate than their domestic counterparts. In contrast to traditional indices, which tend to define agglomeration as the amount of activity taking place in a particular geographic unit, the index constructed in this paper is spatially continuous and thus unbiased with respect to the scale of geographic units and the level of spatial aggregation.

Our analysis presents a rich array of new stylized facts that shed light on the worldwide agglomeration patterns of multinational and domestic firms. We show that the offshore agglomeration patterns of MNCs are distinctively different from those of their headquarters and their domestic counterparts. First, across different types of establishment, multinational headquarters are, on average, the most agglomerative. For example, the average probability of agglomeration at 50 kilometers (km) is 0.8 for MNC headquarters, 0.48% for MNC foreign subsidiaries, and 0.43% for domestic plants. Second, the agglomeration of multinational foreign subsidiaries exhibits a low correlation with the agglomeration of domestic plants, suggesting that the offshore clusters of MNCs are not merely a projection of the domestic clusters. Third, multinational foreign subsidiaries are significantly more agglomerative than domestic plants in capital-, skilled-labor-, and R&D-intensive industries. For example, in industries with above-median capital intensity, the probability of agglomeration at 50 km is, on average, 0.1 percentage point (or equivalently 23%) higher for MNC foreign subsidiaries than for domestic plants.

We then further explore the stylized facts and analyze how different agglomeration economies—including input–output linkages, labor and capital-good market externalities, and technology diffusion—might account for the variations in the agglomeration patterns of MNC and domestic establishments. Our empirical analysis shows that the relative importance of the agglomeration forces varies sharply for MNC offshore subsidiaries, MNC headquarters, and domestic plants. The potential benefits of technology diffusion and capital-good market externality play a significantly stronger role in the agglomeration of MNCs' foreign subsidiaries than in the agglomeration of domestic plants in the same industry. For example, a 10-percentage-point increase in industry-pair technology linkage—measured by the share of patent citations between two industries—increases the probability of agglomeration at 50 km by 0.16 percentage points (or 46%) more for MNC foreign subsidiaries than for domestic plants. Compared to domestic plants and MNC foreign subsidiaries, MNC headquarters' agglomeration patterns are even more strongly influenced by technology diffusion factors. Labor market externality and input–output linkages, in contrast, play a greater role in accounting for the agglomeration patterns of domestic plants.

These findings are largely consistent with the characteristics of multinational firms. Relative to their domestic counterparts in the same industry, MNC offshore subsidiaries are, on average, more knowledge and capital intensive and have stronger motives than domestic plants to agglomerate with each other when their industries exhibit potential for technology diffusion and capital-good market externality. Domestic plants, in contrast, tend to be more concerned about labor-market externality and geographic proximity to input suppliers and customers. Moreover, the increasing segmentation of activities within the boundaries of multinational firms can explain why the agglomeration patterns of MNC foreign subsidiaries differ from those of MNC headquarters. In particular, the input-sourcing focus of offshore production motivates MNC foreign subsidiaries to take into account not only technology diffusion but also capital-good market externality in their location decisions, while a greater emphasis on knowledge-intensive activities—such as R&D, management, and services—leads MNC headquarters to be more driven by technology diffusion benefits.

Our paper builds on an extensive empirical literature in regional and urban economics that examines the importance of Marshallian agglomeration forces in domestic economic geography. Economic historians and regional and urban economists have long recognized the agglomeration of economic activity as one of the most salient features of economic development.[4] However, relatively few studies have investigated the growing spatial concentrations of multinational production around the world and their patterns and driving forces in comparison to those of domestic firms. An overview of the existing literature is beyond the scope of our paper; we focus below on the empirical studies most closely related to our analysis.[5]

As noted earlier, a central issue in agglomeration studies is the measurement of agglomeration. Ellison and Glaeser's (1997) influential paper introduces a "dartboard" approach to construct an index of spatial concentration. The authors note that even in an industry with no tendency for clustering, random locations may not generate regular location patterns due to the fact that the number of plants is never arbitrarily large. Their index thus 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 that had been noted to affect the accuracy of previous indices. Using this index, Rosenthal and Strange (2001) evaluate the importance of agglomeration forces in explaining the localization of U.S. industries and find that both labor-market pooling and input–output linkages have a positive impact on U.S. agglomeration. Overman and Puga (2010), also using Ellison and Glaeser's (1997) index, examine the role of labor-market pooling and input sharing in determining the spatial concentration of UK manufacturing establishments. They find that sectors whose establishments experience more idiosyncratic employment volatility and 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 (see Section 2.2 for a detailed description). Applying this index, Ellison et al. (2010) (henceforth "EGK") employ an innovative empirical approach that exploits the coagglomeration of U.S. industries to disentangle the effects of Marshallian agglomeration economies. Like Rosenthal and Strange (2001), they find a particularly important role for input–output relationships.

Exploring the role of agglomeration economies in MNCs' location patterns also relates our paper to a literature in international trade assessing MNCs' agglomeration decisions. Several studies (see, for example, Head et al., 1995; Head and Mayer, 2004a; Bobonis and Shatz, 2007; Debaere et al., 2010) have examined the role of distance and production linkages in individual multinationals' location decisions. The results of these studies, which suggest that MNCs with vertical linkages tend to agglomerate within a host country/region, shed light on the role of vertical production relationships in the economic geography of multinational production.

Our analysis, assessing the different patterns underlying the global agglomeration of multinational and non-multinational firms, contributes to the literature in several ways. First, instead of examining domestic agglomeration patterns in an individual country, we offer a perspective on the structure of industrial agglomeration around the world. Second, we investigate how the agglomeration of the most

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

[5] Another important strand of empirical literature concerns one of the key theoretical predictions of New Economic Geography models: factor prices should vary systematically across locations with respect to market access. See, for example, Redding and Venables (2004) and Hanson (2005) for related empirical evidence. Among the latest contributors to this literature are Ahlfeldt et al. (2012), who introduce a structural estimation approach incorporating both location fundamentals and agglomeration economies. The authors combine a quantitative model of city structure with the natural experiment of Berlin's division and reunification and find that the model accounts for the observed changes in factor prices and employment.

mobile and distinctive group of firms—the multinationals—compare to the agglomeration of domestic firms. Third, we evaluate how agglomeration economies, particularly the value of external scale economies in knowledge and capital goods, affect MNCs relative to domestic firms, given MNCs' vertically-integrated organizational form and large investment in technologies and capital goods. While existing studies have offered evidence of agglomeration economies in domestic economic geography, little is known about how their influence on the global economic geography of multinationals differs from their influence on the economic geography of domestic firms. Fourth, we examine micro-agglomeration patterns by constructing and exploring plant-level agglomeration indices. Specifically, we examine how a given plant's characteristics—such as size, age, foreign ownership, and the number of products—and its industry's characteristics—such as capital-, skilled-labor-, and R&D-intensity—might jointly explain the extent of agglomeration centered around the plant.

The rest of the paper is organized as follows. Section 2 describes the data and the methodology with which we quantify the agglomeration of multinational and domestic firms and the agglomeration economies driving them. Section 3 presents the stylized facts emerging from the worldwide agglomeration patterns of multinational and domestic firms. Section 4 reports the empirical analysis that assesses the relative importance of agglomeration economies in the agglomeration of MNCs and domestic firms. The last section concludes.

2. Quantifying agglomeration patterns and economies: data and methodology

In this section, we describe the data and the empirical methodology we use to quantify the global agglomeration of multinational and domestic firms and the economic factors that could systematically account for the observed agglomeration patterns.

2.1. The WorldBase database

Our empirical analysis uses a unique worldwide establishment dataset, WorldBase, that covers more than 43 million public and private establishments in more than 100 countries and territories. WorldBase is compiled by Dun & Bradstreet (D&B), a leading source of commercial credit and marketing information since 1845. D&B—presently operating in over a dozen countries either directly or through affiliates, agents, and associated business partners—compiles data from a wide range of sources including public registries, partner firms, telephone directory records, and websites.[6] All information collected by D&B is verified centrally via a variety of manual and automated checks.[7]

2.1.1. Cross-country coverage and geocode information

D&B's WorldBase is, in our view, an ideal data source for the research question proposed in this study. It offers several 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. Examples of cross-border clusters include the metal-working and electrical-engineering cluster involving Germany and German-speaking Switzerland; an electric-machinery cluster involving Switzerland and Italy; a biotech cluster spreading across Germany, Switzerland, and France; an automobile industry cluster that crosses the border of Germany and Slovakia; the Ontario–Canada–Michigan–US (Windsor–Detroit) auto cluster; and the Texas–Northeastern-Mexico cluster. Our data shows that more than 20% of MNC establishment pairs that are within 200 km of each other are in two different countries. The percentage rises to 40% at 400 km. This is not surprising given countries' growing participation in regional trading blocs and the rapid declines in cross-border trade costs.

Second, the database reports detailed information for multinational and domestic, offshore and headquarters establishments. This makes it possible to compare agglomeration patterns across different types of establishment and to investigate how the economic geography of production varies with the organization form of the firm.

Third, the WorldBase database reports the physical address and postal code of each plant, whereas most existing datasets report business registration addresses. The physical location information enables us to obtain precise latitude and longitude information for each plant in the data and compute the distance between each establishment pair. Existing studies have tended to use distance between administrative units, such as state distances, as a proxy for distance of establishments. In doing so, the establishments proximate in actual distance but separated by administrative boundaries (for example, San Diego and Phoenix) can be considered dispersed. Conversely, the establishments far apart but still in the same administrative unit (for example, 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.[8]

2.1.2. MNC and domestic establishment data

Our empirical analysis is based on MNC offshore subsidiaries, MNC headquarters, and domestic plants in 2005. WorldBase reports, for each establishment in the dataset, detailed information on location, ownership, and activities. Four categories of information are used in this paper: (i) industry information including the four-digit SIC code of the primary industry in which each establishment operates; (ii) ownership information including headquarters, domestic parent, global parent, status (for example, joint venture and partnership), and position in the hierarchy (for example, branch, division, and headquarters); (iii) detailed location information for both establishment and headquarters; and (iv) operational information including sales, employment, and year started.

An establishment is deemed 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

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

[7] Early uses of D&B data include, for example, Lipsey's (1978) comparisons of the D&B data with existing sources with regard to the reliability of U.S. data. More recently, Harrison et al. (2004) use D&B's cross-country foreign ownership information. Other research that has used D&B data includes Rosenthal and Strange's (2003) analysis of micro-level agglomeration in the United States; Acemoglu et al.'s (2009) cross-country study of concentration and vertical integration; Alfaro and Charlton's (2009) analysis of vertical and horizontal activities of multinationals; and Alfaro and Chen's (2012) study of the response of multinational firms to the recent global financial crisis.

[8] To account for other forms of trade barriers, such as border, language, and tariffs, we also estimated a measure of trade cost between each pair of plants based on conventional gravity-equation estimations. The trade cost information was then used to construct the index of agglomeration following the empirical methodology described in the next subsection. Alternatively, we computed the agglomeration index based on distance by assuming country borders to have an infinite effect on trade cost. This essentially excluded all establishment pairs located in two different countries, regardless of their actual distance, and focused exclusively on establishments located in the same country. See the HBS working paper version (#10-043) for more detail.

another establishment.[9] We drop establishments with zero or missing employment values and industries with fewer than 10 observations.[10]

Our final sample includes 32,427 MNC offshore manufacturing plants. Top industries include 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, the United Kingdom, Canada, France, Poland, the Czech Republic, and Mexico.

To examine the coverage of our MNC establishment data, we compared U.S. owned subsidiaries in the WorldBase database with the U.S. Bureau of Economic Analysis' (BEA) Direct Investment Abroad Benchmark Survey, a legally mandated confidential survey conducted every five years that covers virtually the entire population of U.S. MNCs. The comparison revealed similar accounts of establishments and activities between the two databases. We also compared WorldBase with UNCTAD's Multinational Corporation Database. These two databases differ in that the former reports at the plant level and the latter at the firm level. For the U.S. and other major FDI source countries, the two databases report similar numbers of firms, but WorldBase contains more plants. See Alfaro and Charlton (2009) for a detailed discussion of the WorldBase data and comparisons with other data sources.

2.2. Quantifying agglomeration patterns

As noted in Head and Mayer's (2004b) study, the measurement of agglomeration is a central challenge in the economic geography literature. There has been a continuous effort to design an index that accurately reflects the agglomeration of economic activities. One of the latest advances in this literature is that of Duranton and Overman (2005) who construct an index to measure the significance of agglomeration in the U.K. DO's index has been adapted by other studies such as EGK who examine the U.S. industries' coagglomeration patterns. We extend this index to assess and compare the agglomeration of multinational and domestic firms worldwide.

The empirical procedure to construct the extended agglomeration index consists of three steps. In the first step, we estimate a distance density function for each pair of industries (including within- and between-industry pairs) based on the distance between MNC establishments. In the second step, we obtain counterfactual density functions based on domestic manufacturing plants in the same industry pair to control for location fundamental factors that affect the location decisions of both domestic and multinational plants. In the last step, we construct the MNC agglomeration index to measure the extent to which multinational establishments in an industry pair are more or less likely to agglomerate than the domestic counterfactuals at a given threshold distance. We repeat the procedure for MNC foreign subsidiaries, MNC foreign subsidiaries weighted by workers, and MNC headquarters.

2.2.1. Step 1: MNC distance density functions

We first estimate MNC's distance 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 (as is the case with the data we use, as described above), distance is only an approximation of the true transport cost between establishments. One source of systematic error, for example, is that the travel time for any given distance might differ between low- and high-density areas. Given the potential noise in the measurement of transport cost, we follow DO in adopting kernel smoothing when estimating the distance density function.

Let τ_{ij}^{M} denote the distance between MNC establishment i and j. For each industry pair k and $\widetilde{k}$, we obtain a kernel density estimator at any level of distance τ (i.e., $f_{k\widetilde{k}}^{M}(\tau)$):

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

where n_k^M and $n_{\widetilde{k}}^M$ are the numbers of MNC establishments 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. This step generates an estimated distance probability density function for each of the 8001 manufacturing industry pairs in our data.

In addition to estimating the distance density functions based on individual establishments, 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 density estimator by weighing each establishment by employment size, given by

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

where r_i^M and r_i^M represent the numbers of employees in MNC establishments i and j, respectively.

2.2.2. Step 2: domestic counterfactual density functions

In the second step, we obtain counterfactual distance density functions based on domestic plants in the same industry pair. By using domestic plants in the same industries as the counterfactuals, the procedure controls for location fundamental factors that affect the location decisions of both MNC and domestic plants. It also enables us to compare the agglomeration patterns of MNC and domestic plants and examine how the agglomeration economies might affect them differently.

Let τ_{ij}^{D} denote the distance between domestic establishments i and j. For each industry pair k and $\widetilde{k}$, we obtain a kernel density estimator at any level of distance τ (i.e., $f_{k\widetilde{k}}^{D}(\tau)$):

$$f_{k\widetilde{k}}^{D}(\tau)=\frac{1}{n_k^D n_{\widetilde{k}}^D h}\sum_{i=1}^{n_k^D}\sum_{j=1}^{n_{\widetilde{k}}^D}K\left(\frac{\tau-\tau_{ij}^D}{h}\right), \tag{3}$$

where n_k^D and $n_{\widetilde{k}}^D$ are the numbers of domestic plants in industries k and $\widetilde{k}$.

Alternatively, we obtain a weighted kernel density estimator for domestic plants by weighing each domestic establishment by employment size:

$$f_{w,k\widetilde{k}}^{D}(\tau)=\frac{1}{h\sum_{i=1}^{n_k^D}\sum_{j=1}^{n_{\widetilde{k}}^D}\left(r_i^D r_j^D\right)}\sum_{i=1}^{n_k^D}\sum_{j=1}^{n_{\widetilde{k}}^D}r_i^D r_j^D K\left(\frac{\tau-\tau_{ij}^D}{h}\right) \tag{4}$$

where r_i^D and r_i^D represent the numbers of employees in domestic establishments i and j, respectively.

[9] 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 paper, 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 (2003) for theoretical work and Chen (2011) for supportive empirical evidence in this area. One can use a methodology similar to the one outlined in the next sub-section to study intra-firm interaction (see Duranton and Overman, 2008).

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

2.2.3. Step 3: MNC agglomeration indices

Next we construct the MNC agglomeration indices using domestic plants as the benchmark. For each industry pair k and $\tilde{k}$, we obtain

$$agglomeration^{M}_{k\tilde{k}}(T) \equiv \sum_{\tau=0}^{T}\left[f^{M}_{k\tilde{k}}(\tau)-f^{D}_{k\tilde{k}}(\tau)\right] \quad (5)$$

or employment-weighted

$$agglomeration^{M}_{w,k\tilde{k}}(T) \equiv \sum_{\tau=0}^{T}\left[f^{M}_{w,k\tilde{k}}(\tau)-f^{D}_{w,k\tilde{k}}(\tau)\right]. \quad (6)$$

Note that $\sum_{\tau=0}^{T} f^{M}_{k\tilde{k}}(\tau)$ and $\sum_{\tau=0}^{T} f^{M}_{w,k\tilde{k}}(\tau)$, the sum of distance density from $\tau=0$ to $\tau=T$, capture the probability of MNC establishments in a given industry pair agglomerating with one another within a threshold distance T. Similarly, $\sum_{\tau=0}^{T} f^{D}_{k\tilde{k}}(\tau)$ and $\sum_{\tau=0}^{T} f^{D}_{w,k\tilde{k}}(\tau)$, the sum of distance density for domestic plants, capture the probability of domestic plants in the same industry pair agglomerating with one another within the same threshold distance. The MNC agglomeration indices $agglomeration^{M}_{k\tilde{k}}(T)$ and $agglomeration^{M}_{w,k\tilde{k}}(T)$ thus are essentially MNCs' differences from domestic establishments in the probabilities of agglomeration and measure the extent to which MNC establishments are more or less likely to agglomerate than their domestic counterfactuals. We compute the index at various distance thresholds, including 50, 100, 200, 400 and 800 km (including thresholds previously considered by DO and EGK as well as lower levels such as 50 and 100 km).

In addition to the pairwise-industry agglomeration index, we also follow the above procedure and construct an agglomeration density measure for each MNC and domestic establishment to measure the probability that a plant is proximate to other plants (from either the same or other industries). The plant-level agglomeration measure enables us to explore the patterns of agglomeration at the micro-plant level and examine how plant characteristics—such as MNC ownership—and industry attributes might jointly explain the different levels of agglomeration observed across plants.

Our methodology to calculate the MNC agglomeration indices, extended based on Duranton and Overman (2005), addresses two key issues that arise with traditional measures of agglomeration, most of which equalize agglomeration with activities located in the same administrative or geographic region (measured by number of firms or volume of production in the region). First, the traditional measures often cannot separate the geographic concentration of the manufacturing industry due to location attractiveness from agglomeration. Second, previous measures, by equating agglomeration with activities in the same region, can omit agglomerating activities separated by administrative or geographic borders, while overestimating the degree of agglomeration within the same administrative or geographic units. The accuracy of these measures is thus dependent on the scale of geographic units. Ellison and Glaeser (1997) develop an index that solves the first problem. DO address the remaining issue of the dependence of existing measures on the level of geographic disaggregation by developing a continuous-space concentration index.

The MNC agglomeration indices thus exhibit three important properties essential to agglomeration measures. First, it is comparable across industries and establishments and captures cross-industry or cross-establishment variation in the level of agglomeration. Second, its construction is based on a counterfactual approach and controls for the effect of location factors—such as market size, natural resources, and policies—that apply to establishments in the same industry. Third, by taking into account spatial continuity, the index is unbiased with respect to the scale and aggregation of geographic units.

However, this methodology also poses two constraints. First, the index requires detailed physical location information for each establishment. As described above, the WorldBase dataset, supplemented by a geocoding software, satisfies this requirement. Second, the empirical procedure to construct the index can be extremely computationally intensive, especially for large datasets. Constructing the index for different types of establishment further increases the computational burden. Given that measuring the agglomeration of all domestic manufacturing plants worldwide is infeasible with the size of the WorldBase dataset and the computational intensity of the empirical procedure, we adopt a random sampling strategy as EGK. For each SIC 3-digit industry with more than 1000 observations, we obtain a random sample of 1000 plants. For industries with fewer than 1000 observations, we include all domestic plants. This yields a final sample of 127,897 domestically owned plants and 32,427 MNC offshore manufacturing plants.

2.3. Measuring agglomeration economies

We now turn to economic factors that could systematically account for the observed agglomeration patterns of MNC and domestic plants. Four categories of agglomeration economies have been stressed in the literature of economic geography, including: (i) vertical production linkages, (ii) externality in labor markets, (iii) externality in capital-good markets, and (iv) technology diffusion.[11] However, the advantage of geographic proximity and subsequently the importance of agglomeration economies can differ dramatically between multinational and domestic firms and between MNC foreign subsidiaries and headquarters. For instance, given their technology intensity, MNCs can find technology diffusion from other MNCs in closely linked industries particularly attractive and thus have greater incentives to agglomerate with other MNCs that share close technology linkages. We discuss below the role of each agglomeration economy in multinational firms' location choices and the proxies used to represent each force.

2.3.1. 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 agglomeration incentive also applies to MNCs, given their large volumes of sales and intermediate inputs.[12] 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 (for example, Dell and Intel; Ford and Delphi) can be particularly strong.[13] However, MNCs, on the other hand, engage in substantial intra-firm trade, sourcing a significant share of their inputs within the boundary of the firm. This distinctive organization structure suggests that compared to domestic firms, the location decisions of MNC establishments could also be less driven by external input–output relationships.

To determine the importance of customer and supplier relationships in multinationals' vs. domestic plants' agglomeration decisions, we construct a variable, $IOlinkage_{k\tilde{k}}$, to measure the extent of the input–output relationship between each pair of industries. We use the 2002 Benchmark Input–Output Data (specifically, the Detailed-level Make, Use and Direct Requirement Tables) published by the Bureau of Economic

[11] In addition to agglomeration economies, the location fundamentals of multinational production—such as country market size, comparative advantage, and trade cost—also affect the location decisions of multinational firms. In the paper, we use worldwide domestic establishment locations as the counterfactual to account for the role of location fundamentals. In a robustness analysis, we also constructed an expected index of agglomeration, reflecting the geographic distribution of MNC plants predicted exclusively by country- and region-level location factors of multinational production, including, for example, market size, trade costs, comparative advantage, infrastructure, corporate taxes (see the HBS working paper version (#10-043) for more detail).

[12] For FDI theoretical literature in this area, see, for example, Krugman (1991), Venables (1996), and Markusen and Venables (2000).

[13] Head et al. (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.

Analysis, and define $IOlinkage_{k\tilde{k}}$ as the share of industry k's inputs that come directly from industry $\tilde{k}$ and vice versa. These shares are calculated relative to all input–output flows including those of non-manufacturing industries and final consumers. Table A.1 reports the summary statistics of industry-level control variables. 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, which, as shown in Table A.2, is highly correlated. We used the mean values in our analysis, but obtained similar results when we used the maximum measure.

2.3.2. 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.[14] Externalities can also occur as workers move from one job to another, especially between firms characterized by similar skill requirements.[15]

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 (such as 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, $labor_{k\tilde{k}}$, using the correlation in occupational percentages.

2.3.3. 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. Geographically concentrated industries offer better support to providers of capital goods (such as producers of specialized components and providers of machinery maintenance) and reduce their risk of investment (due, for example, to the existence of resale markets).[16] Local expansion of capital-intensive activities can consequently lead to expansion of the supply of capital goods, thereby reducing the cost of capital goods.

To evaluate the role of capital-good market externalities, we construct a new measure of industries' similarity in capital-good demand —in a spirit similar to the measure of industries' similarity in labor demand—using capital flow data from the Bureau of Economic Analysis (BEA). The capital flow table (CFT), a supplement to the 1997 Benchmark Input–Output (I–O) accounts, shows detailed purchases of capital goods (such as motors and generators, textile machinery, mining machinery and equipment, wood containers and pallets, computer storage devices, and wireless communications equipment) by using industry. We compute—for each using industry—the share of investment in each capital good and then measure each industry pair's similarity in capital-good investment, denoted by $capitalgood_{k\tilde{k}}$, using the industry pair's correlation in investment shares.[17]

2.3.4. Technology diffusion

A fourth motive relates to the diffusion of technologies. Technology can diffuse from one firm to another through movement of workers, 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. 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, for example, Jaffe et al., 2000; EGK), using patent citation flow data taken from the NBER Patent Database. The data, compiled by Hall et al. (2001), includes 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 (such as technology classification and citations of prior art) and about the inventors submitting the application (such as name and city). We construct the technology diffusion variable, that is, $technology_{k\tilde{k}}$, by measuring the extent to which technologies in industry k cite technologies in industry $\tilde{k}$, and vice versa.[18] In practice, there is little directional difference in $technology_{k\tilde{k}}$ due to the extensive number of citations within a single technology field. We obtain both maximum and mean for each set of pairwise industries. We used the mean values in our analysis, but obtained similar results when using the maximum measure.

Constructing the proxies of agglomeration economies using the U.S. industry-level account data is 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. 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 those of most other countries, enabling us to dissect linkages between disaggregated product categories.

Table A.2 presents the correlation matrix. As shown, the proxies of agglomeration economies have very low correlations. For example, the correlation between industry-pair input–output linkage and similarity in capital-good demand is about 0.19 and the correlation between production linkage and technology diffusion is 0.29. This suggests that industry pairs exhibit significant variation in their relatedness in inputs, labor, capital goods, and technology. For example, industry pairs with strong input–output linkages often have weak linkages in capital goods and technology. This provides us a key source of variation for disentangling the effects of agglomeration economies.

[14] This argument has been formally considered in Marshall (1890), Krugman (1991), and Helsley and Strange (1990). Rotemberg and Saloner (2000), for a related motivation, argue that workers can benefit because multiple firms offer protection against ex-post appropriation of investments in human capital.

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

[16] Agglomeration can also create costs, for example, by increasing labor and capital-good prices. Like benefits, these costs can be 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.

[17] 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.

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

Table 1
Descriptive statistics for MNC and domestic agglomeration densities.

	Obs.	Mean	Std. dev.	Min.	Max.
MNC foreign subsidiaries					
Threshold (T) = 50 km	8001	0.005	0.001	0.001	0.013
T = 100 km	8001	0.009	0.002	0.002	0.025
T = 200 km	8001	0.019	0.005	0.005	0.051
T = 400 km	8001	0.041	0.010	0.011	0.112
T = 800 km	8001	0.097	0.024	0.025	0.254
MNC foreign subsidiaries (employment-weighted)					
Threshold (T) = 50 km	8001	0.004	0.003	0.000	0.023
T = 100 km	8001	0.008	0.005	0.000	0.042
T = 200 km	8001	0.017	0.009	0.000	0.077
T = 400 km	8001	0.036	0.019	0.000	0.128
T = 800 km	8001	0.079	0.037	0.001	0.243
Domestic plants					
Threshold (T) = 50 km	8001	0.004	0.002	0.000	0.048
T = 100 km	8001	0.008	0.003	0.001	0.091
T = 200 km	8001	0.019	0.007	0.003	0.191
T = 400 km	8001	0.045	0.015	0.010	0.379
T = 800 km	8001	0.101	0.032	0.029	0.472
Domestic plants (employment-weighted)					
Threshold (T) = 50 km	8001	0.003	0.002	0.000	0.046
T = 100 km	8001	0.007	0.003	0.001	0.089
T = 200 km	8001	0.015	0.007	0.003	0.178
T = 400 km	8001	0.037	0.015	0.007	0.331
T = 800 km	8001	0.087	0.031	0.021	0.458
MNC headquarters					
Threshold (T) = 50 km	8001	0.008	0.002	0.003	0.031
T = 100 km	8001	0.016	0.004	0.005	0.059
T = 200 km	8001	0.032	0.008	0.009	0.117
T = 400 km	8001	0.069	0.018	0.020	0.251
T = 800 km	8001	0.153	0.043	0.043	0.544

Notes: This table reports the summary statistics of agglomeration densities for MNC foreign subsidiaries, domestic plants, and MNC headquarters at 50, 100, 200, 400 and 800 km. All industry pairs (SIC3) are included.

Table 2
Within- and between-industry agglomeration densities.

	Within-industry		Between-industry	
	Obs.	Mean	Obs.	Mean
MNC foreign subsidiaries				
Threshold (T) = 50 km	126	0.006	7875	0.005
T = 100 km	126	0.011	7875	0.009
T = 200 km	126	0.022	7875	0.019
T = 400 km	126	0.048	7875	0.041
T = 800 km	126	0.108	7875	0.096
MNC foreign subsidiary (employment-weighted)				
Threshold (T) = 50 km	126	0.006	7875	0.004
T = 100 km	126	0.012	7875	0.008
T = 200 km	126	0.024	7875	0.015
T = 400 km	126	0.049	7875	0.036
T = 800 km	126	0.104	7875	0.079
Domestic plants				
Threshold (T) = 50 km	126	0.007	7875	0.004
T = 100 km	126	0.014	7875	0.008
T = 200 km	126	0.029	7875	0.019
T = 400 km	126	0.065	7875	0.045
T = 800 km	126	0.135	7875	0.107
Domestic plants (employment-weighted)				
Threshold (T) = 50 km	126	0.006	7875	0.003
T = 100 km	126	0.013	7875	0.007
T = 200 km	126	0.026	7875	0.015
T = 400 km	126	0.057	7875	0.036
T = 800 km	126	0.116	7875	0.087
MNC headquarters				
Threshold (T) = 50 km	126	0.009	7875	0.008
T = 100 km	126	0.018	7875	0.015
T = 200 km	126	0.037	7875	0.032
T = 400 km	126	0.078	7875	0.069
T = 800 km	126	0.171	7875	0.153

Notes: This table reports the summary statistics of agglomeration densities for within- and between-industry pairs, respectively, for MNC foreign subsidiaries, domestic plants, and MNC headquarters at 50, 100, 200, 400 and 800 km.

3. The global agglomeration of MNCs and domestic plants: stylized facts

In this section, we examine the global agglomeration patterns of MNC and domestic plants and present emerging stylized facts.

First, we show in Table 1 the descriptive statistics of MNC and domestic plants' (cumulative) agglomeration densities at various threshold distances (i.e., $\sum_{\tau=0}^{T} f^{M}_{\widetilde{kk}}(\tau)$, $\sum_{\tau=0}^{T} f^{M}_{w,\widetilde{kk}}(\tau)$, $\sum_{\tau=0}^{T} f^{D}_{\widetilde{kk}}(\tau)$ and $\sum_{\tau=0}^{T} f^{D}_{w,\widetilde{kk}}(\tau)$, the probability of MNC/domestic establishments agglomerating at a threshold distance). We find that multinational headquarters exhibit, on average, the highest probability of agglomeration among the different types of establishment. At 50 km, for example, the average probability of agglomeration is 0.8% for MNC headquarters, 0.48% for MNC foreign subsidiaries, and 0.43% for domestic plants. At 100 km, the average probability of agglomeration increases to 1.6% for MNC headquarters, 0.92% for MNC foreign subsidiaries, and 0.85% for domestic plants. The differences between MNC headquarters and the other types of establishment are statistically significant at all threshold distances, while the differences between MNC foreign subsidiaries and domestic plants are statistically significant at 50 and 100 km.

The above finding is summarized as our first stylized fact below.

Stylized fact 1: *Across different types of establishment, multinational headquarters are, on average, the most agglomerative.*

Stylized fact 1 is broadly consistent with the knowledge capital theory of multinational firms (see Markusen, 2002), which predicts that MNC headquarters should concentrate in skilled-labor-abundant countries and subsidiaries should be relatively dispersedly distributed across host regions based on markets and comparative advantages. Our finding also lends empirical support to theoretical predictions in urban economics which suggest greater clustering of headquarters relative to that of manufacturing plants (see, for example, Duranton and Puga, 2005).

Among MNC foreign subsidiaries, industry pairs in which MNCs exhibit some of the highest offshore agglomeration probabilities include, as reported in Table A.3, Footwear except Rubber (314) and Boot and Shoe Cut Stock and Findings (313); Knitting Mills (225) and Footwear except Rubber (314); Dolls, Toys, Games (394) and Sporting and Athletic and Footwear except Rubber (314); Miscellaneous Publishing (274) and Paperboard Mills (263); and Miscellaneous Publishing (274) and Miscellaneous Transportation Equipment (379).

In Table 2, we present descriptive statistics for within- and between-industry pairs, respectively. We find that (i) Stylized fact 1 holds for both within- and between-industry pairs[19]; and (ii) establishments in the same industry are significantly more agglomerative than establishments from different industries. The latter observation is consistent with the expectation noted in EGK that firms from different industry pairs exhibit greater variation in their relatedness in production, factor markets, and technology space, thereby displaying weaker average agglomeration incentives.

[19] In the descriptive and the empirical analyses, we take into account both within- and between-industry agglomeration. As a robustness, we also analyzed only between-industry agglomeration, also called “coagglomeration”, and obtained similar results. As noted by EGK, 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. While location fundamentals and all agglomeration economies tend to predict spatial concentration among firms in the same industry, their predictions of which industry pairs should agglomerate vary significantly. Between-industry agglomeration patterns thus offer an important source of variation for separating the effects of different agglomeration economies.

Table 3
Correlations of MNC and domestic agglomeration densities.

	50 km (sub)	100 km (sub)	200 km (sub)	400 km (sub)	800 km (sub)	50 km (dom)	100 km (dom)	200 km (dom)	400 km (dom)	800 km (dom)
MNC foreign subsidiaries vs. domestic plants										
T = 50 km (sub)	1.00									
T = 100 km (sub)	0.99	1.00								
T = 200 km (sub)	0.99	0.99	1.00							
T = 400 km (sub)	0.96	0.97	0.99	1.00						
T = 800 km (sub)	0.89	0.92	0.95	0.98	1.00					
T = 50 km (dom)	**0.17**	0.17	0.16	0.15	0.12	1.00				
T = 100 km (dom)	0.18	**0.18**	0.17	0.16	0.13	0.99	1.00			
T = 200 km (dom)	0.20	0.19	**0.19**	0.18	0.15	0.99	0.99	1.00		
T = 400 km (dom)	0.24	0.24	0.24	**0.23**	0.21	0.96	0.97	0.98	1.00	
T = 800 km (dom)	0.31	0.31	0.31	0.31	**0.29**	0.79	0.81	0.85	0.92	1.00
MNC foreign subsidiaries vs. domestic plants (employment weighted)										
T = 50 km (sub)	1.00									
T = 100 km (sub)	0.99	1.00								
T = 200 km (sub)	0.99	0.99	1.00							
T = 400 km (sub)	0.99	0.99	0.99	1.00						
T = 800 km (sub)	0.95	0.95	0.96	0.98	1.00					
T = 50 km (dom)	**0.33**	0.33	0.33	0.33	0.32	1.00				
T = 100 km (dom)	0.33	**0.33**	0.33	0.33	0.32	0.99	1.00			
T = 200 km (dom)	0.33	0.33	**0.33**	0.33	0.32	0.99	0.99	1.00		
T = 400 km (dom)	0.33	0.33	0.33	**0.34**	0.33	0.96	0.97	0.98	1.00	
T = 800 km (dom)	0.31	0.31	0.32	0.32	**0.33**	0.81	0.82	0.86	0.93	1.00

	50 km (sub)	100 km (sub)	200 km (sub)	400 km (sub)	800 km (sub)	50 km (hq)	100 km (hq)	200 km (hq)	400 km (hq)	800 km (hq)
MNC foreign subsidiaries vs. MNC headquarters										
T = 50 km (sub)	1.00									
T = 100 km (sub)	0.99	1.00								
T = 200 km (sub)	0.99	0.99	1.00							
T = 400 km (sub)	0.96	0.97	0.99	1.00						
T = 800 km (sub)	0.90	0.92	0.95	0.98	1.00					
T = 50 km (hq)	**0.44**	0.46	0.48	0.52	0.56	1.00				
T = 100 km (hq)	0.43	**0.45**	0.48	0.52	0.57	0.99	1.00			
T = 200 km (hq)	0.41	0.43	**0.46**	0.52	0.58	0.99	0.99	1.00		
T = 400 km (hq)	0.37	0.39	0.44	**0.51**	0.59	0.96	0.97	0.99	1.00	
T = 800 km (hq)	0.35	0.38	0.43	0.51	**0.59**	0.92	0.94	0.97	0.99	1.00

Notes: This table reports the correlations between the agglomeration densities of MNC foreign subsidiaries and domestic plants and between MNC foreign subsidiaries and MNC headquarters at 50, 100, 200, 400 and 800 km. The bold emphasis corresponds to indices at the same threshold distance.

Next, we examine in Table 3 the correlations of agglomeration density measures across different types of plant. Comparing the probability of agglomeration between MNC foreign subsidiaries and domestic plants, we find the correlation of the two to be 0.17 at 50 km (and only slightly higher at more aggregate distance levels), suggesting that multinational and non-multinational plants exhibit sharply different spatial patterns. In over 65% of the industry pairs, MNC foreign subsidiaries are more likely to agglomerate than domestic plants. The agglomeration patterns of MNC headquarters and foreign subsidiaries are correlated with a higher coefficient of 0.44 at 50 km, implying that while for some industry pairs the clusters of MNC subsidiaries resemble those of headquarters, for other industry pairs the two types of establishment exhibit distinctly different agglomeration patterns.

These observations, summarized in Stylized fact 2, indicate that the offshore clusters of MNCs are not merely a projection of the domestic clusters. The driving forces of MNCs' offshore agglomeration are likely to vary from those of domestic plants and MNC headquarters, as we explore in Section 4.[20]

Stylized fact 2: *The agglomeration of multinational foreign subsidiaries exhibits a low correlation with the agglomeration of domestic plants.*

Next we take a first glance at the differences in multinational and domestic plants' agglomeration patterns and examine how they relate to industry characteristics such as capital intensity, skilled-labor intensity, and R&D intensity.[21]

In Fig. 1, we plot the distributions of pairwise industries' agglomeration densities at 50 km for multinational foreign subsidiaries and domestic plants, respectively. We find that for industries with greater than median levels of capital intensity, the distribution shifts rightward for multinational foreign subsidiaries compared with domestic plants. The probability of agglomeration in these relatively capital-intensive industries is, on average, 0.1 percentage point (or equivalently 23% of the mean value) higher for MNC foreign subsidiaries than for domestic plants. This pattern is similarly observed for industries with greater than median levels of skilled-labor intensity and R&D intensity: in skilled-labor- and R&D-intensive industries, the distribution of multinational foreign subsidiaries' agglomeration densities dominates the distribution of domestic plants, with the mean difference of around 0.1 percentage point.[22]

Table 4 presents the Kolmogorov–Smirnov (K–S) first-order stochastic dominance test on the distributions of MNC subsidiaries and domestic plants in capital-, skilled-labor- and R&D-intensive industries.

[20] Similarly, the correlations do not change significantly when we drop within-industry pairs (which consist of 126 observations).

[21] We use the NBER-CES Manufacturing Industry Database to construct each industry's capital and skilled-labor intensities, which are defined as, respectively, the ratios of investment and of non-production workers' payroll to value added. Each industry's R&D intensity is measured using the median firm's ratio of R&D expenditure relative to value added based on the Compustat database.

[22] The pattern, again, does not change when within-industry agglomeration indices are excluded.

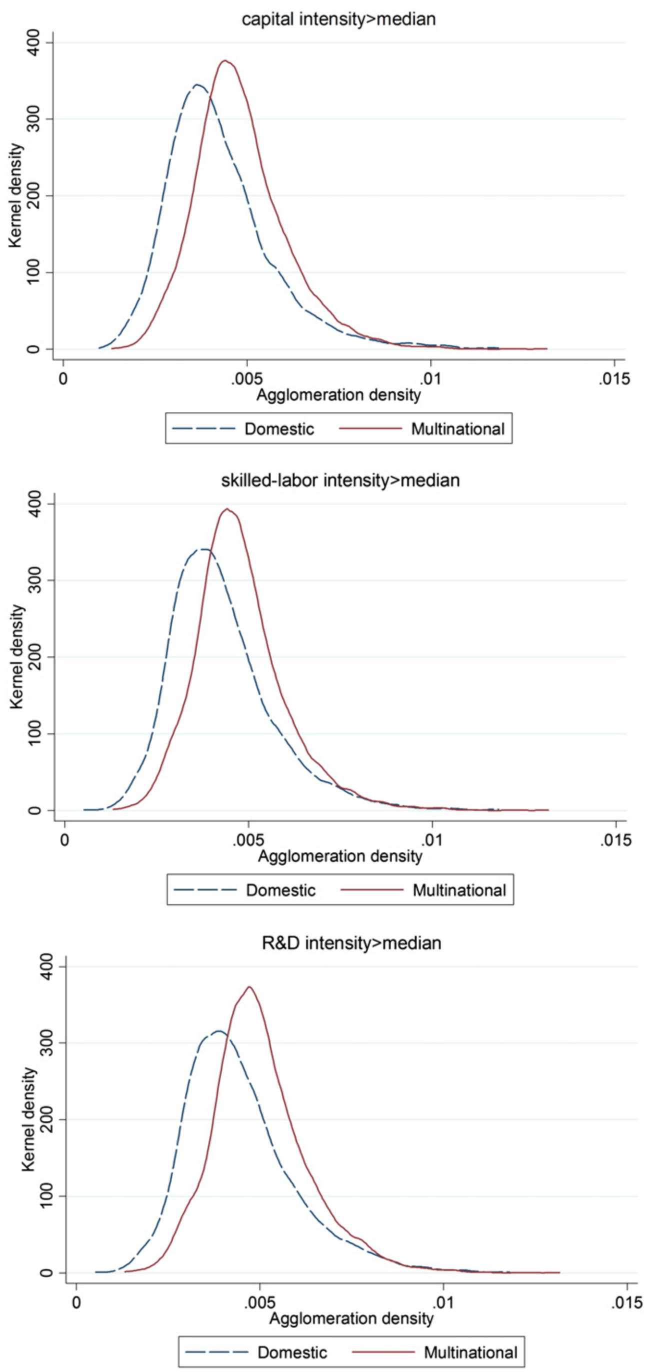

Fig. 1. The agglomeration-density distributions of multinational foreign subsidiaries and domestic plants: pairwise industry level.

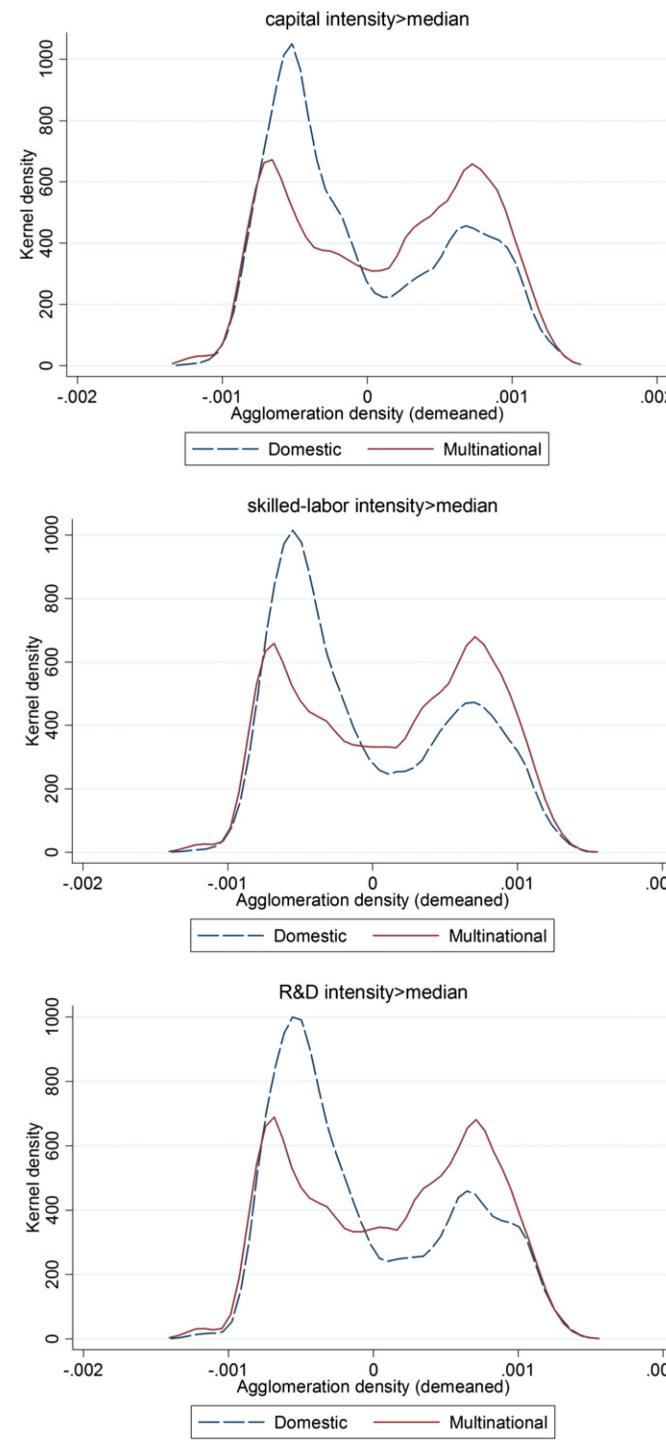

Fig. 2. The agglomeration-density distributions of multinational foreign subsidiaries and domestic plants: plant level.

Table 4
Comparing the distance density distributions of MNC and domestic plants: first-order stochastic dominance test.

Industry characteristics	Two-sided K–S test p-value	One-sided K–S test p-value
	MNC foreign subsidiary = domestic	MNC foreign subsidiary < domestic
Capital intensity > median	0.00	0.00
Skilled-labor intensity > median	0.00	0.00
R&D intensity > median	0.00	0.00

Notes: This table reports the Kolmogorov–Smirnov test results that compare MNC and domestic plants' distance density distributions at 50 km for relatively capital-, skilled-labor-, and R–D-intensive industries. "Capital intensity" is the industry's ratio of capital expenditure relative to value added; "skilled-labor intensity" is the ratio of non-production workers' payroll to value added. Both variables are computed based on the NBER-CES Manufacturing Industry Database. "R&D intensity" is the median firm's ratio of R&D expenditures relative to value added in an industry, computed based on the Compustat database.

Table 5
The roles of industry and plant characteristics in plant-level agglomeration.

Dependent variable	T = 50 km	T = 50 km	T = 100 km	T = 100 km
Plant agglom. index				
MNC dummy	−0.001***	−0.001***	−0.003***	−0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
× IO linkages	0.002	0.003	0.006	0.006
	(0.002)	(0.002)	(0.004)	(0.004)
× capital intensity	0.005***	0.005***	0.009***	0.009***
	(0.002)	(0.002)	(0.004)	(0.004)
× skilled-labor intensity	0.002***	0.002***	0.004***	0.004***
	(0.000)	(0.000)	(0.001)	(0.001)
× R&D intensity	0.0003*	0.0003*	0.001*	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Revenue		0.001***		0.001***
		(0.000)		(0.000)
Age		0.00004**		0.0001**
		(0.000)		(0.000)
Product count		0.000		0.000
		(0.000)		(0.000)
State–industry FE	Yes	Yes	Yes	Yes
Obs.	122,426	122,324	122,426	122,324
R2	0.997	0.997	0.997	0.997

Notes: This table estimates the roles of industry and plant characteristics in explaining the plant-level agglomeration densities at 50, 100, 200, 400 and 800 km. Bootstrapped standard errors are reported in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1. All regressions include region–industry fixed effect. "MNC dummy" is an indicator of MNC foreign subsidiaries. "Revenue", "age" and "product count" are a plant's logged revenue, age, and number of products, respectively. "Capital intensity" is the industry's ratio of capital expenditure relative to value added; "skilled-labor intensity" is the ratio of non-production workers' payroll to value added. Both variables are computed based on the NBER-CES Manufacturing Industry Database. "R&D intensity" is the median firm's ratio of R&D expenditures relative to value added in an industry, computed based on the Compustat database.

We find that the differences of the distributions are statistically significant and, further, the one-sided test suggests that MNC subsidiaries are significantly more agglomerative than domestic plants in capital-, skilled-labor- and R&D-intensive industries.

We also plot the distribution of agglomeration densities at the plant level, for multinational foreign subsidiaries and domestic plants, respectively. As discussed in Section 2.2, we also compute—in addition to pairwise-industry agglomeration measures—a plant-level agglomeration measure to capture the degree to which a plant is proximate to other plants. The plant-level densities are demeaned by industry averages to ensure within-industry comparisons. Similar to industry-level patterns, we show in Fig. 2 that multinational foreign subsidiaries exhibit greater agglomeration than their domestic peers in capital-, skilled-labor-, and R&D-intensive industries.[23]

In Table 5, we examine how plant characteristics—such as ownership structure, size, age and the number of products—and industry characteristics—including capital, skilled-labor, and R&D intensity—might jointly explain the extent of agglomeration centered around each plant. The estimation results at 50 and 100 km are reported. To control for the role of location fundamentals, a vector of region-industry dummies is also included in the analysis.

We find that the degree of agglomeration varies sharply across plants in the same region and industry. First, multinational foreign subsidiaries are significantly more agglomerative than domestic plants in the same capital-, skilled-labor-, and R&D-intensive industries. This result suggests that multinational foreign subsidiaries enjoy greater agglomeration benefits than their domestic counterparts do when industrial activities are capital- and knowledge-intensive.[24] Second, plant size also matters. At both 50 and 100 km, we find that plants with a larger revenue are significantly more likely to attract agglomeration. This is similarly true for older plants. On the other hand, the number of products produced by each plant does not appear to have a significant effect on agglomeration.

These findings, summarized as our Stylized fact 3 below, suggest that in industries with high capital, skilled-labor, and R&D requirements, MNCs—which tend to be more productive and more capital- and knowledge-intensive than domestic plants in the same industry—are more likely to provide as well as derive benefits of capital market externality and technology diffusion—than their domestic peers—and thus are more likely to cluster with each other offshore.

Stylized fact 3: *Multinational foreign subsidiaries are more agglomerative than domestic plants in capital-, skilled-labor-, and R&D-intensive industries.*

4. Assessing the roles of agglomeration economies

Next we further explore the stylized facts and examine how different agglomeration economies—including input–output linkages, labor and capital-good market externalities, and technology diffusion—might account for the variations in the agglomeration patterns of MNC and domestic establishments.

4.1. MNC offshore agglomeration

We first consider MNCs' offshore agglomeration. Specifically, we evaluate how agglomeration economies affect the agglomeration patterns of MNC foreign subsidiaries relative to their domestic counterfactuals by estimating the following equation:

$$agglomeration^M_{k\tilde{k}}(T) \equiv \sum_{\tau=0}^{T}\left[f^M_{k\tilde{k}}(\tau)-f^D_{k\tilde{k}}(\tau)\right] = \beta_1 IOlinkage_{k\tilde{k}} + \beta_2 capitalgood_{k\tilde{k}} + \beta_3 labor_{k\tilde{k}} + \beta_4 technology_{k\tilde{k}} + \varepsilon_{ij}, \quad (7)$$

where $agglomeration^M_{k\tilde{k}}(T)$, capturing the differences between MNC foreign subsidiaries and domestic plants in their probabilities of agglomeration, measures the extent to which MNC establishments are more or less likely to agglomerate than their domestic counterfactuals, and the coefficients β_1, β_2, β_3, and β_4 represent the differences in the effects of the covariates on multinational foreign subsidiaries and domestic plants.

Table 6 reports the regression results. We find that the proxy for technology diffusion exerts a stronger effect on MNC foreign subsidiaries than on domestic plants in the same industry pairs. For example, at 50 km a 10-percentage-point increase in the level of technology diffusion—that is, the share of patent citations between two industries—raises the agglomeration probability of MNC foreign subsidiaries by 0.03 percentage points (or 7%) more than the agglomeration probability of domestic plants. Industry pairs' correlations in capital-good demand, a proxy for potential capital-good market externality, also exert a stronger effect on the agglomeration of MNC foreign subsidiaries at 400 km and above. Interestingly, the proxy for potential labor-market externality, captured by industry-pair correlations in labor demand, has a greater effect on the agglomeration of domestic plants than on the agglomeration of MNC foreign subsidiaries.[25]

These results are consistent with the stylized facts documented in Section 3 and suggest that, given the technology- and capital-intensive characteristics of multinational firms, it is important to take into account technology diffusion and capital-good market externality in explaining MNCs' offshore agglomeration.

[23] This result is similarly confirmed by the two-sided and the one-sided Kolmogorov–Smirnov (K–S) first-order dominance tests.

[24] We also considered including a separate dummy variable to represent MNCs' domestic subsidiaries and found that the agglomeration patterns of MNC domestic subsidiaries are fairly similar to that of domestic plants in the MNC headquarters country when controlling for plant characteristics. The result is available upon request.

[25] We also examined distance thresholds lower than 50 km, such as 20 km which is less than the distance between, say, the JFK airport and Newark. The results were quantitatively similar to those at 50 km.

Table 6
The roles of agglomeration economies in the agglomeration of MNC foreign subsidiaries vs. domestic plants.

Dependent variable MNC subsidiary agglomeration index	T = 50 km	T = 100 km	T = 200 km	T = 400 km	T = 800 km
IO linkages	−0.001	−0.001	−0.002	−0.005	−0.010
	(0.001)	(0.001)	(0.003)	(0.006)	(0.009)
Capital good	−0.000	−0.000	0.000	0.001*	0.003**
	(0.000)	(0.000)	(0.000)	0.000	(0.001)
Labor	−0.0002**	−0.0004**	−0.001**	−0.002**	−0.003**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Technology	0.003*	0.006*	0.013*	0.025*	0.042**
	(0.002)	(0.004)	(0.008)	(0.015)	(0.022)
Obs.	8001	8001	8001	8001	8001
R2	0.61	0.61	0.62	0.67	0.76
IO linkages	−0.007	−0.007	−0.007	−0.007	−0.007
Capital good	−0.004	−0.003	0.001	0.010	0.021
Labor	−0.026	−0.026	−0.027	−0.026	−0.019
Technology	0.022	0.021	0.021	0.019	0.016

Notes: This table estimates the role of agglomeration economies in explaining how MNC foreign subsidiaries agglomerate relative to domestic counterfactuals in the same industry at 50, 100, 200, 400 and 800 km. Bootstrapped standard errors are reported in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All regressions include industry fixed effect. “IO linkages” is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. “Capital goods” is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). “Labor” is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry-Occupation Employment Matrix. “Technology” measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data. Normalized beta coefficients are shown in the lower panel.

Table 7
The roles of agglomeration economies in the employment-weighted agglomeration of MNC foreign subsidiaries vs. domestic plants.

Dependent variable MNC subsidiary weighted agglomeration index	T = 50 km	T = 100 km	T = 200 km	T = 400 km	T = 800 km
IO linkages	−0.002*	−0.005*	−0.01*	−0.017	−0.016
	(0.001)	(0.003)	(0.006)	(0.012)	(0.018)
Capital good	0.0002	0.0005	0.001*	0.003**	0.009***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.003)
Labor	−0.0000	−0.0001	−0.0002	−0.001	−0.0002
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
Technology	0.016***	0.029***	0.056***	0.107***	0.198***
	(0.004)	(0.007)	(0.014)	(0.026)	(0.046)
Industry FE	Yes	Yes	Yes	Yes	Yes
Obs.	8001	8001	8001	8001	8001
R2	0.37	0.37	0.37	0.40	0.48
IO linkages	−0.023	−0.023	−0.023	−0.019	−0.009
Capital good	0.020	0.020	0.023	0.031	0.047
Labor	−0.001	−0.003	−0.006	−0.008	−0.001
Technology	0.073	0.073	0.071	0.068	0.063

Notes: This table estimates the role of agglomeration economies in explaining how employment-weighted MNC foreign subsidiaries agglomerate relative to domestic counterfactuals in the same industry at 50, 100, 200, 400 and 800 km. Bootstrapped standard errors are reported in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All regressions include industry fixed effect. “IO linkages” is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. “Capital goods” is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). “Labor” is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry-Occupation Employment Matrix. “Technology” measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data. Normalized beta coefficients are shown in the lower panel.

The lower panel of Table 6 reports the normalized beta coefficients.[26] Comparing the standardized coefficients of agglomeration forces, we find that the effect of technology diffusion outweighs the effect of capital-good market correlations. This implies that technology diffusion benefits are likely to be a more important factor than capital-good market externality in MNCs' offshore agglomeration decisions, especially at the disaggregated distance levels. Comparing the magnitudes of the normalized beta coefficients across distance thresholds, we find that the impact of technology diffusion diminishes and the effect of capital-good market externality rises at more aggregate distance levels. The stronger effect of technology diffusion at lower distance levels suggests that, compared to the other agglomeration economies, benefits from technology diffusion tend to be localized geographically.

Thus far, we have examined MNC offshore agglomeration using the 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 MNC foreign subsidiary 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.

The results are reported in Table 7. Similarly, proxies for technology diffusion and capital-good market externality exert a stronger effect on MNC foreign subsidiaries than on domestic plants in the same industry pairs. For example, at 50 km a 10-percentage-point increase in the level of technology diffusion—the share of patent citations between two industries—leads to a 0.16-percentage-point (or 46%) greater increase in the agglomeration probability of MNC foreign subsidiaries than the agglomeration probability of domestic plants. Again, technology diffusion, an agglomeration force that involves close labor interaction and

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

Table 8
The roles of agglomeration economies in the agglomeration of MNC headquarters vs. domestic plants.

Dependent variable MNC HQ agglom. index	T = 50 km	T = 100 km	T = 200 km	T = 400 km	T = 800 km
IO linkages	−0.002	−0.004	−0.007	−0.011	−0.008
	(0.001)	(0.003)	(0.005)	(0.010)	(0.014)
Capital good	−0.0001	−0.0002	−0.0004	−0.0003	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Labor	−0.0003**	−0.001**	−0.001**	−0.003***	−0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)
Technology	0.004***	0.008***	0.016***	0.036***	0.082***
	(0.001)	(0.003)	(0.005)	(0.011)	(0.022)
Industry FE	Yes	Yes	Yes	Yes	Yes
Obs.	8001	8001	8001	8001	8001
R2	0.71	0.72	0.74	0.80	0.87
IO linkages	−0.019	−0.018	−0.016	−0.011	−0.004
Capital good	−0.011	−0.011	−0.008	−0.002	0.005
Labor	−0.034	−0.034	−0.034	−0.029	−0.023
Technology	0.032	0.031	0.031	0.031	0.034

Notes: This table estimates the role of agglomeration economies in explaining how MNC headquarters agglomerate relative to domestic counterfactuals in the same industry at 50, 100, 200, 400 and 800 km. Bootstrapped standard errors are reported in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All regressions include industry fixed effect. "IO linkages" is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. "Capital goods" is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). "Labor" is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry-Occupation Employment Matrix. "Technology" measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data. Normalized beta coefficients are shown in the lower panel.

mobility, appears to be the strongest agglomeration factor at all distance thresholds. The role of the input–output relationship is now significantly stronger for domestic plants than for MNC foreign subsidiaries at disaggregated distance levels, but is not significantly different at more aggregate distance levels (such as 400 km and 800 km). The effect of labor-market externality is not significantly different between MNC foreign subsidiaries and domestic plants in this case.

The above findings are consistent with the characteristics of multinational firms. Relative to their domestic counterparts, multinationals exhibit greater participation in capital- and technology-intensive activities. As a result, in industries with strong potential for capital-good market externality and technology diffusion, MNCs are more likely to realize these agglomeration economies when they agglomerate with other, productive and capital- and knowledge-intensive MNCs. In contrast, domestic plants—with lower capital- and technology-intensity—place a greater emphasis on proximity to local suppliers and customers and thus have greater incentives to agglomerate with upstream and downstream industries.

4.2. MNC headquarter agglomeration

We next examine the patterns of MNC headquarter agglomeration relative to the agglomeration of domestic plants. Table 8 reports the estimation results. Technology diffusion exerts a significantly stronger effect on MNC headquarters than on domestic plants in the same industry pairs. In contrast, potential labor market externality exerts, again, a stronger effect on the agglomeration of domestic plants.

Comparing the normalized beta coefficients in Table 8 with those in Table 6, we find that (i) technology diffusion exerts a stronger effect on MNCs' headquarter agglomeration than their agglomeration overseas, and (ii) capital-good market externality exerts a stronger

Table 9
The roles of agglomeration economies in the process of MNC agglomeration.

Dependent variable MNC subsidiary agglomeration index	T = 50 km	T = 100 km	T = 200 km	T = 400 km	T = 800 km
IO linkages	−0.002	−0.004	−0.007	−0.011	−0.008
	(0.001)	(0.003)	(0.005)	(0.010)	(0.014)
Capital good	−0.0001	−0.0002	−0.0004	−0.0003	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Labor	−0.0003**	−0.001**	−0.001**	−0.003***	−0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)
Technology	0.004***	0.008***	0.016***	0.036***	0.082***
	(0.001)	(0.003)	(0.005)	(0.011)	(0.022)
Industry FE	Yes	Yes	Yes	Yes	Yes
Obs.	8001	8001	8001	8001	8001
R2	0.71	0.72	0.74	0.80	0.87
IO linkages	−0.019	−0.018	−0.016	−0.011	−0.004
Capital good	−0.011	−0.011	−0.008	−0.002	0.005
Labor	−0.034	−0.034	−0.034	−0.029	−0.023
Technology	0.032	0.031	0.031	0.031	0.034

Notes: This table estimates the role of agglomeration economies in explaining how new MNC foreign subsidiaries agglomerate with existing MNC foreign subsidiaries in comparison to their domestic counterfactuals in the same industry at 50, 100, 200, 400 and 800 km. Bootstrapped standard errors are reported in parentheses, ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. All regressions include industry fixed effect. "IO Linkages" is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. "Capital goods" is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). "Labor" is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry-Occupation Employment Matrix. "Technology" measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data. Normalized beta coefficients are shown in the lower panel.

effect on MNCs' offshore agglomeration than on the agglomeration of MNC headquarters. These results suggest that the agglomeration of MNC foreign subsidiaries, with their input-sourcing focuses, is more influenced by capital-good market externality, whereas the agglomeration of headquarters, with their specialization in providing services such as R&D and management, is more driven by technology diffusion benefits.

4.3. The process of MNC agglomeration

To shed light on the formation of MNC clusters, in particular, the spatial interdependence between incumbents and entrants, we now turn from the geographic patterns to the process of multinational agglomeration. Doing so also helps us to address the different establishment dates of plants. Our estimates thus far take into account not only new plants' entry decisions but also incumbents' decisions to continue in their current locations. But the mix of old and new plants could create the potential for reverse causality between MNC location patterns and measures of agglomeration economies.[27]

We therefore explore in this subsection the dynamics of location decisions. Specifically, we distinguish new from incumbent plants and assess new MNC plants' propensity to agglomerate with incumbents. Repeating the procedure described in Section 2.2, we construct an index of agglomeration between MNC entrants in 2004–2005 and MNC incumbents established before 2004 and use domestic plants as the benchmark. For each industry pair k and $\widetilde{k}$, the index measures the extent to which new MNC subsidiaries in industry k are more likely to cluster with incumbent MNCs in industry $\widetilde{k}$ and vice versa relative to their domestic counterfactuals.[28]

Table 9 reports the estimates. The roles of agglomeration forces remain robust in explaining the entry patterns of MNCs. Relative to domestic plants, multinational entrants display a stronger propensity to cluster with incumbent multinationals when technology diffusion benefits are greater. Labor-market externality, again, has a stronger impact on the agglomeration of domestic plants.

5. Conclusion

The emergence of multinational clusters is one of the most notable phenomena in the process of globalization. In this paper, we examine the global patterns and forces of MNC agglomeration—both offshore and at headquarters—relative to the patterns and forces of domestic-firm agglomeration. Our analysis, using a worldwide plant-level dataset and a novel index of agglomeration, yields a number of new insights into the industrial landscape of multinational production.

First, the offshore clusters of MNCs are not simply a reflection of domestic industrial clusters. Across different types of plant, multinational headquarters are, on average, the most agglomerative, followed by multinational foreign subsidiaries and domestic plants. Further, the agglomeration densities of MNC foreign subsidiaries, MNC headquarters, and domestic plants exhibit only limited correlations, suggesting that multinationals follow distinctively different agglomeration patterns offshore than their domestic counterparts do. In particular, multinational foreign subsidiaries are significantly more agglomerative than domestic plants in capital-, skilled-labor-, and R&D-intensive industries.

Second, exploring the patterns of the multinational agglomeration, we find that multinationals' location choices are significantly influenced by agglomeration economies including technology diffusion and capital-good market externality. The impact of technology diffusion, in particular, outweighs the effect of all other agglomeration economies.

Third, the importance of agglomeration economies varies sharply between MNCs' offshore agglomeration and the agglomeration of MNC headquarters and domestic plants. MNCs' offshore plants are significantly more influenced than domestic plants by technology diffusion and capital-good market externality factors.

One potential extension of our analysis that is worthy of particular attention is to explore how patterns of MNC agglomeration vary across regions. For example, labor market externality can offer a stronger incentive for agglomeration in countries with more rigid and less mobile labor markets. Similarly, the varying quality of infrastructure across regions can affect the value of proximity for vertically linked industries. Firms are likely to have a stronger motive to cluster with suppliers and customers when they are in a country with poorer infrastructure. Further analysis of the role of regional characteristics in determining the clustering of MNCs could yield additional insights.

Appendix A

Table A.1
Descriptive statistics for agglomeration economies.

	Mean	Std. dev.	Min.	Max.
Input–output (IO) linkages	0.003	0.012	0.000	0.193
Capital good	0.476	0.209	0.004	1.000
Labor	0.333	0.227	0.014	1.000
Technology	0.007	0.012	0.000	0.179

Notes: This table reports the summary statistics of agglomeration economy measures. "IO linkages" is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. "Capital goods" is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). "Labor" is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry–Occupation Employment Matrix. "Technology" measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data.

Table A.2
Correlations 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: This table reports the correlations between agglomeration economy measures. "IO linkages" is the average of an industry pair's shares of inputs that come directly from each other, calculated using the U.S. Benchmark Input–Output Data. "Capital goods" is an industry pair's correlation in their investments in capital goods, calculated based on the BEA capital flow table (CFT). "Labor" is an industry pair's correlation in their use of occupations, calculated using the BLS's National Industry–Occupation Employment Matrix. "Technology" measures the extent to which an industry pair's patents cite each other, constructed using the USPTO data.

[27] We also examined regional agglomeration patterns from which the United States is excluded to alleviate concerns of endogenous agglomeration economy measures and found the results to be robust. 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. See the HBS working paper version (#10-043) for more details.

[28] To address the possibility that the index of MNC agglomeration might reflect the clustering between MNC and domestic plants, we also used an alternative benchmark, the agglomeration between new MNC subsidiaries and incumbent domestic plants. We find that for each industry pair, new MNC foreign subsidiaries exhibit a stronger tendency to agglomerate with incumbent MNC plants than with incumbent domestic plants. Moreover, the estimated effects of the agglomeration economies remain largely similar.

Table A.3
Top industry pairs by the MNC foreign-subsidiary agglomeration index.

MNC foreign-subsidiary agglomeration index			
	T = 200 km		
274	Miscellaneous publishing	379	Miscellaneous transportation equipment
314	Footwear, except rubber	313	Boot and shoe cut stock and findings
225	Knitting mills	313	Boot and shoe cut stock and findings
367	Electronic components and accessories	225	Knitting mills
225	Knitting mills	314	Footwear, except rubber
	T = 400 km		
274	Miscellaneous publishing	379	Miscellaneous transportation equipment
314	Footwear, except rubber	313	Boot and shoe cut stock and findings
225	Knitting mills	313	Boot and shoe cut stock and findings
274	Miscellaneous publishing	213	Chewing and smoking tobacco and snuff
263	Paperboard mills	213	Chewing and smoking tobacco and snuff
Employment-weighted MNC foreign-subsidiary agglomeration index			
	T = 200 km		
394	Dolls, toys, games and sporting	314	Footwear, except rubber
394	Dolls, toys, games and sporting	313	Boot and shoe cut stock and findings
225	Knitting mills	314	Footwear, except rubber
314	Footwear, except rubber	313	Boot and shoe cut stock and findings
225	Knitting mills	394	Dolls, toys, games and sporting and athletic
	T = 400 km		
394	Dolls, toys, games and sporting	314	Footwear, except rubber
394	Dolls, toys, games and sporting	313	Boot and shoe cut stock and findings
225	Knitting mills	314	Footwear, except rubber
314	Footwear, except rubber	313	Boot and shoe cut stock and findings
225	Knitting mills	313	Boot and shoe cut stock and findings

Notes: Authors' calculations.

References

Acemoglu, Daron, Johnson, Simon, Mitton, Todd, 2009. Determinants of vertical integration: financial development and contracting costs. J. Financ. 63, 1251–1290.
Ahlfeldt, Gabriel, Redding, Stephen J., Sturn, Daniel, Wolf, Nikolaus, 2012. The economics of density: evidence from the Berlin Wall. CEP Discussion Papers dp1154. Centre for Economic Performance, LSE.
Alfaro, Laura, Charlton, Andrew, 2009. Intra-industry foreign direct investment. Am. Econ. Rev. 99 (5), 2096–2119.
Alfaro, Laura, Chen, Maggie Xiaoyang, 2012. Surviving the global financial crisis: foreign ownership and establishment performance. Am. Econ. J.: Econ. Pol 4 (3), 30–55.
Antràs, Pol, Helpman, Elhanan, 2004. Global sourcing. J. Polit. Econ. 112 (3), 552–580.
Antràs, Pol, Helpman, Elhanan, 2008. Contractual frictions and global sourcing. In: Helpman, Elhanan, Marin, Dalia, Verdier, Thierry (Eds.), The Organization of Firms in a Global Economy. Harvard University Press, Cambridge, pp. 9–54.
Bobonis, Gustavo J., Shatz, Howard J., 2007. Agglomeration, adjustment, and state policies in the location of foreign direct investment in the United States. Rev. Econ. Stat. 89 (1), 30–43.
Chen, Maggie, 2011. Interdependence in multinational production networks. Can. J. Econ. 44 (3), 930–956.
Debaere, Peter, Lee, Joonhyung, Paik, Myungho, 2010. Agglomeration, backward and forward linkages: evidence from South Korean investment in China. Can. J. Econ. 43 (2), 520–546.
Duranton, Gilles, Overman, Henry, 2005. Testing for localization using microgeographic data. Rev. Econ. Stud. 72 (4), 1077–1106.
Duranton, Gilles, Overman, Henry, 2008. Exploring the detailed location patterns of U.K. manufacturing industries using microgeographic data. J. Reg. Sci. 48 (1), 213–243.
Duranton, Gilles, Puga, Diego, 2004. Micro-foundations of urban agglomeration economies. In: Henderson, J.Vernon, Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics vol. 4. Elsevier, Amsterdam: North-Holland, pp. 2063–2117.
Duranton, Gilles, Puga, Diego, 2005. From sectoral to functional urban specialisation. J. Urban Econ. 57 (2), 343–370.
Ellison, Glenn, Glaeser, Edward, 1997. Geographic concentration in U.S. manufacturing industries: a dartboard approach. J. Polit. Econ. 105 (5), 889–927.
Ellison, Glenn, Glaeser, Edward, Kerr, William, 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. Am. Econ. Rev. 100 (3), 1195–1213.
Hall, Bronwyn H., Jaffe, Adam B., Trajtenberg, Manuel, 2001. The NBER patent citation data file: lessons, insights and methodological tools. National Bureau of Economic Research Working Paper 8498.
Hanson, Gordon, 2005. Market potential, increasing returns and geographic concentration. J. Int. Econ. 67 (1), 1–24.
Harrison, Ann E., Love, Inessa, McMillan, Margaret S., 2004. Global capital flows and financing constraints. J. Dev. Econ. 75 (1), 269–301.
Head, Keith, Mayer, Thierry, 2004a. Market potential and the location of Japanese firms in the European Union. Rev. Econ. Stat. 86 (4), 959–972.
Head, Keith, Mayer, Thierry, 2004b. The empirics of agglomeration and trade. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics vol. 4. Elsevier, Amsterdam, pp. 2609–2669.
Head, Keith, Ries, John, Swenson, Deborah, 1995. Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States. J. Int. Econ. 38 (3–4), 223–247.
Helpman, Elhanan, Melitz, Marc, Yeaple, Stephen, 2004. Export versus FDI with heterogeneous firms. Am. Econ. Rev. 94 (1), 300–316.
Helsley, Robert W., Strange, William C., 1990. Matching and agglomeration economies in a system of cities. Reg. Sci. Urban Econ. 20 (2), 189–212.
Jaffe, Adam, Trajtenberg, Manuel, Fogarty, Michael, 2000. Knowledge spillovers and patent citations: evidence from a survey of inventors. Am. Econ. Rev. Pap. Proc. 90 (2), 215–218.
Krugman, Paul, 1991. Increasing returns and economic geography. J. Polit. Econ. 99 (3), 483–499.
Lipsey, Robert E., 1978. The creation of microdata sets for enterprises and establishments. Ann. INSEE 30 (31), 395–422.
Markusen, James, 2002. Multinational Firms and the Theory of International Trade. MIT Press, Cambridge.
Markusen, James, Venables, Anthony J., 2000. The theory of endowment, intra-industry and multi-national trade. J. Int. Econ. 52 (2), 209–234.
Marshall, Alfred, 1890. Principles of Economics. MacMillan and Co., London.
Navaretti, Giorgio Barba, Venables, Anthony J., 2006. Multinational Firms in the World Economy. Princeton University Press.
Ottaviano, Gianmarco I.P., Puga, Diego, 1998. Agglomeration in the global economy: a survey of the "new economic geography". World Econ. 21 (6), 707–731.
Ottaviano, Gianmarco, Thisse, Jacques-François, 2004. Agglomeration and economic geography. In: Henderson, J. Vernon, Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics vol. 4. Elsevier, Amsterdam, pp. 2563–2608.
Overman, Henry G., Puga, Diego, 2010. Labour Pooling as a Source of Agglomeration: An Empirical Investigation. In: Economics, Agglomeration, Glaeser, Edward L. (Eds.), Chicago. University of Chicago Press, IL, pp. 133–150.
Puga, Diego, 2010. The magnitude and causes of agglomeration economies. J. Reg. Sci. 50 (1), 203–219.
Redding, Stephen, 2010. The empirics of new economic geography. J. Reg. Sci. 50 (1), 297–311.
Redding, Stephen, 2011. Economic geography: a review of the theoretical and empirical literature. The Palgrave Handbook of International Trade, (Chapter 16).
Redding, Stephen, Venables, Anthony, 2004. Economic geography and international inequality. J. Int. Econ. 62 (1), 53–82.
Rosenthal, Stuart, Strange, William, 2001. The determinants of agglomeration. J. Urban Econ. 50 (2), 191–229.
Rosenthal, Stuart, Strange, William, 2003. Geography, industrial organization, and agglomeration. Rev. Econ. Stat. 85 (2), 377–393.
Rosenthal, Stuart, 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.
Rotemberg, Julio, Saloner, Garth, 2000. Competition and human capital accumulation: a theory of interregional specialization and trade. Reg. Sci. 30 (4), 373–404.
Venables, Anthony, 1996. Equilibrium locations of vertically linked industries. Int. Econ. Rev. 37 (2), 341–359.
Yeaple, Stephen, 2003. The complex integration strategies of multinational firms and cross-country dependencies in the structure of foreign direct investment. J. Int. Econ. 60 (2), 293–314.