CESifo WORKING PAPERS

7104
2018

June 2018

The Hardships of Long Distance Relationships: Time Zone Proximity and Knowledge Transmission within Multinational Firms

Dany Bahar

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7104*
Category 8: Trade Policy

The Hardships of Long Distance Relationships: Time Zone Proximity and Knowledge Transmission Within Multinational Firms

Abstract

Using a unique dataset on worldwide multinational corporations with precise location of headquarters and affiliates, I present evidence of a trade-off between distance to the headquarters and the knowledge intensity of the foreign subsidiary's economic activity, emerging from dynamics related to the proximity-concentration hypothesis. This trade-off is strongly diminished the higher the overlap in working hours between the headquarters and its foreign subsidiary. In order to rule out biases arising from confounding factors, I implement a regression discontinuity framework to show that the economic activity of a foreign subsidiary located just across the time zone line that increases the overlap in working hours with its headquarters is, on average, about one percent higher in the knowledge intensity scale. I find no evidence of the knowledge intensity and distance trade-off weakening when a non-stop flight exists between the headquarters and the foreign subsidiary. The findings suggest that lower barriers to real-time communication within the multinational corporation play an important role in the location strategies of multinational corporations.

JEL-Codes: F230, L220, L250.

Keywords: multinational firms, multinational corporations, knowledge, location, proximity concentration hypothesis, FDI.

Dany Bahar
The Brookings Institution
Harvard Center for International Development
Harvard University / Cambridge / MA / USA
db21@post.harvard.edu
http://www.danybahar.com

May 30, 2018
Previously circulated under the title "Heavier than air? Knowledge Diffusion within the Multinational Firm". This paper benefited from helpful comments at various stages from Martin Abel, Laura Alfaro, Pol Antras, Sam Asher, Sebastian Bustos, Marcus Casey, Matias Cattaneo, Michele Coscia, Ricardo Hausmann, Elhanan Helpman, Bill Kerr, Marc Melitz, Paul Novosad, Nathan Nunn, Lant Pritchett, Krista Rasmussen, Brina Seidel, Ran Shorrer, Ernesto Stein, Stephen Yeaple, Muhammed Yildirim, Richard Zeckhauser, Oren Ziv and the participants of the XIX Dynamics in Economic Growth and International Trade conference, as well as attendants in different seminars at Harvard University, the NBER, the Inter-American Development Bank, Georgetown University and LACEA. Jean-Paul Leidenz Font and Samuel Rosenow provided excellent research assistance. I acknowledge financial support from Harvard Economics Department. All errors are my own. Download newest version: https://www.dropbox.com/s/06cdlvoh8bi2kv0/BAHAR_KnowledgeMNC.pdf?dl=0

1 Introduction

About fifty percent of cross-country income variation is explained by differences in productivity.[1] This begs the question: if productivity-inducing knowledge is available in some places, why isn't it easily transferable to others? Arrow (1969) suggests that the transmission of knowledge is difficult and costly. These difficulties arise because effective knowledge transmission involves human interaction, which cannot be fully replaced with written words, neither be embedded in goods that can be shipped at low costs.[2] A firm, as any other economic agent, faces difficulties when transferring knowledge among different divisions and affiliates, as has been extensively explored in the literature (e.g., Oldenski, 2012; Keller and Yeaple, 2013; Giroud, 2012). The ability of a multinational firm to transfer knowledge to its foreign affiliates, however, should depend on several dimensions that characterize the locations of the firm's headquarters and its subsidiaries. This paper explores the role of time zones in reducing the costs of knowledge transmission within multinational corporations (MNCs).

This paper's contribution to the literature is threefold. First, using a highly detailed establishment-level *worldwide* dataset on MNCs I complement the existing empirical evidence suggesting that costs of knowledge transmission faced by MNCs play a role in the proximity-concentration hypothesis, in particular by creating a trade-off between the level of knowledge intensity of a foreign subsidiary's economic activity and the geographic distance between such subsidiary and its global headquarters. Second, I show that this "knowledge and distance trade-off" significantly weakens, and even disappears, the more overlap in working hours there is between a foreign subsidiary and its global headquarters. Third, to rule out the role of other unobservables explaining my results, I use a regression discontinuity designed to exploit discrete spatial variation in time zones and show that foreign subsidiaries located at roughly the same distance from their headquarters are active in economic activities that significantly differ in knowledge intensity depending on whether they are located in a time zone closer to the headquarters or not.

The trade-off between distance and knowledge intensity cannot be explained using earlier frameworks that looked at the proximity-concentration hypothesis and the fragmentation of MNCs which, implicitly or explicitly, assumed zero marginal cost or costs orthogonal to distance of transferring knowledge between headquarters and subsidiaries (i.e., Helpman, 1984; Markusen, 1984; Brainard, 1993; Markusen et al., 1996; Markusen, 1997; Carr et al., 2001; Helpman et al., 2004; Markusen and Maskus, 2002).[3] This trade-off, however, emerges in any model that incorporates the idea of marginal costs of transferring knowledge increasing in geographic distance. For example, Ramondo and Rodriguez-Clare

[1] e.g., Hall and Jones (1999); Caselli (2005)

[2] Knowledge that resides in human minds is usually referred to as tacit (Polanyi, 1966). Tacit knowledge is information that cannot be easily explained, embedded or written down.

[3] A number of empirical studies have tested the validity of these models' predictions, but there has been little or no emphasis on testing the assumption that knowledge transmission is costless.

(2013) study the gains from openness in a comprehensive model that includes both multinational production and trade, and assume that all multinational production activity entails iceberg type efficiency losses that vary across country pairs. This marginal cost, they assume, is an increasing function of geographical and cultural distance. Ramondo (2014) models multinational production where foreign subsidiaries face a combination of a variable and a fixed cost when relying in technology or knowledge from their headquarters, capturing the idea of limited span of control for the headquarters, thus generating a loss of productivity at the affiliate level. Arkolakis et al. (2013) model an economy in which countries specialize either in innovation or in production. When a country specializes in innovation it implies that its firms open production subsidiaries abroad with a marginal cost that affects their productivity. These marginal costs are meant to capture various impediments that multinationals face when operating in a different economic, legal or social environment, as well as the various costs of technology transfer incurred by multinationals in different production locations. Tintelnot (2014) presents a model of multinational production in which foreign plants engage in variable costs that are determined, in part, by distance between the headquarters and the location of the foreign subsidiary. Keller and Yeaple (2013) provide an explanation to why the marginal cost of knowledge transmission increases with distance by modeling such cost as shipping costs for intermediate goods embedding headquarters services.[4] In their model, shipping of intermediate goods is more prevalent for subsidiaries active in knowledge intensive activities. My study builds on the research by Keller and Yeaple (2013), in further exploring the "distance and knowledge intensity" trade-off emerging from their setting studying the proximity-concentration hypothesis. In particular, a contribution of this study is the finding that this "trade-off" cannot be fully explained by trade costs of intermediate goods within the MNC, but also by a better ability of firms to communicate in real time, as measured by a headquarters and its foreign subsidiary having more overlap in working hours (based in their geographic location).

An interpretation of this finding is that not all knowledge can be fully embedded in intermediate goods. Examples of the type of knowledge that gets transferred from the headquarters to its foreign subsidiary are in the form of management, monitoring, coordination, troubleshooting, etc. which is a form of *tacit* knowledge: it lives on people brains and cannot be fully written down or embedded in a machine. This type of knowledge requires human interaction of some sort for its transmission, as suggested by Arrow (1969). Thus, the loss in efficiency that distant foreign affiliates face as evidenced in the literature is likely also related to the difficulties in transferring *tacit* knowledge (Polanyi, 1962). In fact, the consensus in the existing literature on the economics of knowledge is that the transmission of knowledge is not immediate, and that knowledge diffusion strongly decays with distance. For instance, the paper by Jaffe et al. (1993) was among the first to make this claim, showing that patent

[4]See Irarrazabal et al. (2013) for a similar setting which does not include the knowledge dimension.

citations are more frequent within the same geographic area. Bottazzi and Peri (2003) followed up using European data. Along the same lines, Keller (2002) showed that knowledge spillovers decrease with distance by looking at productivity changes as explained by foreign R&D investment. He documents that the half-life of such spillovers is 1200Km. More recently, Bahar et al. (2014) show that a country is 65% more likely to add a new product to its export basket whenever a geographic neighbor is a successful exporter of the same good, a finding that is attributed to the local character of knowledge diffusion. Bahar and Rapoport (2018), consistently with the idea that the transmission of *tacit* knowledge requires human interaction,show that migrants are an important driver of knowledge diffusion across nations.[5] In the context of MNCs, it has also been shown that more complicated tasks require more time and effort for coordination and monitoring, and this becomes much more difficult at longer distances (e.g., Gumpert, 2015).

The empirical exercise in this paper is based on a sample of about 55000 domestic and 25000 foreign horizontal subsidiaries belonging to over 2000 MNCs from the Worldbase dataset by Dun & Bradstreet.[6] For the most part, I focus my attention on MNCs active in the manufacturing sector that have horizontally expanded into foreign countries. For each one of the foreign subsidiaries in the sample I have information on their physical location and primary economic activity, as defined by the 1987 Standard Industry Classification (SIC).

After computing precise distances between each foreign affiliate and its MNC global headquarters following a geocoding process using Google Maps, I document a negative partial correlation between knowledge intensity and the distance between a headquarters and its foreign subsidiaries (after including a number of controls that would account for other possible explanations). For example, everything else equals, the estimation in the paper implies that an American MNC headquartered in Houston in Texas would locate a meat packing subsidiary –an economic activity with low knowledge intensity levels– in Kabul in Afghanistan (approximately 12000Km away from the headquarters) and a semiconductor plant –an economic activity with high levels of knowledge intensity– in Ireland (at approximately 7000Km of distance from the headquarters). A novel source of variation in this empirical exercise is a industry-level knowledge intensity measure that I construct based worker-level characteristics in each industry, as opposed traditional measures that rely on balance-sheet data. The measures aim to capture the *tacit* knowledge intensity of an economic activity by averaging the accumulated experience and training of the workforce of such industry, using occupational characteristics defined in the O*NET project dataset.[7]

[5]See Keller (2004) for a review of this literature.

[6]The dataset was privately acquired from D&B and is not publicly accessible. It has been previously used in the literature by Lipsey (1978), and more recently by Black and Strahan (2002); Harrison et al. (2004); Acemoglu et al. (2009); Alfaro and Charlton (2009); Alfaro and Chen (2012) and Alfaro et al. (2015).

[7]Previous studies have also used the O*NET database to construct industry level measures. For example, Oldenski (2012) measures the importance of communication with the headquarters and importance of the communication with the customer, for each industry. Costinot et al. (2011) uses O*NET to create an industry level measure of task routine-ness

I then use a number of measures that ease the transmission of knowledge by enhancing communication and monitoring between a MNC's headquarters and its foreign subsidiaries. First, the existence of a non-stop flight between airports located within 100Km of both the headquarters and the foreign subsidiary. Second, the number of overlapping working hours in a business day between the locations of both the headquarters and the foreign subsidiary. I then show both descriptively and using a regression discontinuity design that the "knowledge-intensity and distance" trade-off is weakened when there is a larger overlap in working hours between the foreign subsidiary and the headquarters. Thus, the cost of shipping intermediate goods, which would be just as relevant within the same time zone (because north-south shipping is equally as expensive as east-west shipping), is not enough to explain the fact that MNCs tend to locate their foreign subsidiaries active in knowledge intensive geographically nearby, as suggested by Keller and Yeaple (2013).

Overall, the results in this paper suggests that tacit knowledge plays is an important determinant of the concentration-proximity hypothesis, and in particular, the knowledge and distance trade-off that emerges from it. These findings have larger implications for a number of yet-unanswered questions in economics. For instance, high barriers to knowledge transmission may explain persistent differences in productivity levels between countries and the divergence of their incomes over time (e.g., Pritchett 1997, Hall and Jones 1999), because productivity-inducing knowledge does not diffuse easily.

The rest of the paper is divided as follows. Section 2 describes the dataset and the construction of relevant variables. Section 3 presents descriptive evidence documenting, among other results, the distance and knowledge intensity trade-off that arises from the proximity-concentration hypothesis. Section 4 explores how the variables measuring the ease of communication between a foreign subsidiary and its headquarters (i.e., existence of a non-stop flight and the overlap in working hours) affects the aforementioned trade-off. Section 4 also implements a regression discontinuity design to estimate whether knowledge intensity significantly differs with changes in time zones that facilitate communication with the headquarters. Section 5 concludes and addresses areas for future research regarding the role of tacit knowledge in economic activity.

2 Data and Definitions of Variables

2.1 Worldbase dataset by Dun & Bradstreet

I use the Worldbase dataset by Dun & Bradstreet (acquired in May 2012) as the main data source for the empirical exercise. The dataset has information on more than one hundred million establishments worldwide with data from year 2012. Each establishment is uniquely identified and linked to its global

for 77 sectors. Keller and Yeaple (2013) also present results making use of knowledge intensity variables constructed with O*NET in their Appendix. Autor et al. (2003) use O*NET predecessor, DOT, to construct measures for routine and non-routine tasks.

headquarters (referred to as the "global ultimate"). For this study I focus on foreign plants engaged in manufacturing industries (SIC codes 2000 to 3999) owned by MNCs. As suggested by Caves (1971), an MNC is "an enterprise that controls and manages production establishments – plants – located in at least two countries."[8]

The sample obtained from the dataset includes about 55 thousand domestic and 25 thousand foreign horizontal subsidiaries of MNCs scattered active in the manufacturing sector in over 100 countries, which report to over 2000 MNCs.[9] For the analysis, I will use the reported main SIC code as the only indicator of a plant's economic activity. There are about 450 unique SIC 4-digit codes (in manufacturing) reported by subsidiaries as their main economic activity in the dataset. The sample I use is significantly smaller than the overall Dun & Bradstreet dataset (which originally has over 124 million establishments) for several reasons. First, I only include subsidiaries of multinational corporations: that in itself reduces the sample to be about 9 percent of the original dataset. Second, I only use manufacturing subsidiaries, which is a much smaller share the universe of subsidiaries. I limit the sample to subsidiaries in the manufacturing sector, which significantly reduces the sample further, also to about 9 percent of the subsidiaries that are part of a MNC (as known, most of the MNCs are in the service sector, as shown too by Alfaro and Charlton (2009)). Further, the sample has all the domestic subsidiaries but the foreign ones are limited to only those that can be classified as a foreign horizontal expansion (as I explain in next subsection). Also, from Alfaro and Charlton (2009) we know that this is a small share of all subsidiaries. Hence, we end up with a relatively small sample given the initial size of the dataset.

In order to obtain the precise location of each plant I geocode the dataset using Google Maps Geocoding API to find the exact latitude and longitude of its headquarters and each one of its foreign subsidiaries. With this I computed the exact distance between each headquarters and its foreign subsidiaries. Figure 1 maps the unique locations of all foreign subsidiaries (dots) and headquarters (triangles) in the sample.

[Figure 1 about here.]

For instance, Figure 2 shows the headquarters and subsidiaries of an American car manufacturing multinational firm. The firm, headquartered in the US, has a number of foreign subsidiaries on different continents. The lines originating from the headquarters represent the geographic distance to each subsidiary.

[Figure 2 about here.]

Online Appendix Section B discusses this dataset more in detail.

[8] I exclude MNCs for which 99% of their subsidiaries or employees are in the home country, besides them having plants in two or more countries. This drops a small number of Chinese MNCs with one or two subsidiaries in Hong Kong and the rest in China.

[9] I performed on the dataset an algorithm that would group multinationals not only based on their assigned number, but also on their names when the differences are small (e.g., Sony Corporation vs. Sony Corp).

2.2 Main Variables Definitions

2.2.1 Horizontal Foreign Subsidiary

For the plant-level dataset I define a foreign subsidiary as a horizontal expansion based on its SIC code vis-à-vis all the SIC codes reported by the firm, in all of its domestic subsidiaries in the home country. This resolves the data issues that arise when the economic activity of the headquarters does not necessarily represent the main business of the firm. For instance, in the dataset, the headquarters of a well known worldwide multinational in the cosmetic world is defined under SIC code 6719 ("holding company"). However, many of its domestic subsidiaries are classified under SIC code 2844 (perfumes, cosmetics, and other toiletries), which would be a more natural classification for the firm as a whole. Hence, by limiting the definitions to the global ultimate's SIC category only, horizontal relationships would be underestimated.

In the sample, 32% of subsidiaries are foreign horizontal. Similarly to Alfaro and Charlton (2009), I find that some horizontal foreign affiliates also classify as vertical (both downstream and upstream).[10] I keep these observations in the sample, and allow in every empirical specification for a different constant for these types of subsidiaries by adding the proper dummy variables.

2.2.2 Knowledge Intensity Measures

In order to estimate the knowledge intensity of industries I create a new measure that aims to capture the tacit knowledge intensity for each industry. The measures use data from the Occupational Employment Statistics (OES) from the Bureau of Labor Statistics,[11] and occupational profiles compiled by the Occupational Information Network (O*NET) project.[12] OES breaks down the composition of occupations for each industry code,[13] based on a list of about 800 occupations. These occupations can be linked to occupational profiles generated by O*NET, which includes results from a large number of survey questions on the characteristics of each occupation.

The relevant questions in the survey that capture the learning component of the workers, as mentioned above, are the ones related to experience and training. The exact form of the questions from O*NET are:

[10] I follow Alfaro and Charlton (2009) methodology of using the US input-output tables to define vertical relations. Online Appendix Section B.3 expands on this discussion.

[11] Data from 2011, downloadable from ftp://ftp.bls.gov/pub/special.requests/oes/oesm11in4.zip

[12] O*NET is the successor of the US Department of Labor's Dictionary of Occupational Titles (DOT). I use the O*NET database version 17, downloadable from http://www.onetcenter.org/download/database?d=db_17_0.zip. Costinot et al. (2011) also use O*NET to create an industry level measure of task routineness for 77 sectors. Keller and Yeaple (2013) also present results making use of knowledge intensity variables constructed with O*NET in the Appendix.

[13] I used Pierce and Schott (2012) concordance tables to convert industry codes from NAICS to 1987 SIC. The concordance table is downloadable from http://faculty.som.yale.edu/peterschott/files/research/ data/appendix_files_20111004.zip.

- How much related experience (in months) would be required to be hired to perform this job?
- How much "on-site" or "in-plant" training (in months) would be required to be hired to perform this job?
- How much "on-the-job" training (in months) would be required to be hired to perform this job?

Using these questions I generate the main knowledge intensity measure that I will be using in the empirical analysis section. The measure, which I refer to it as *"Experience plus training"* throughout the paper, is constructed by measuring the (wage-weighted) average months of experience plus on-site and on-the-job training required to work in each industry. In particular, for each sector s knowledge intensity is defined as:

$$KI_i = \sum_o \omega_{o,i} cumexp_o$$

Where $cumexp_o$ is the sum of experience and training associated to occupation o, $\omega_{i,o}$ is the weight of occupation o in industry i, measured by either share of employment or wage. Naturally, we have that $\sum_o \omega_{o,i} = 1$ for every i. While I use within-sector wage share to define $\omega_{o,i}$, the results are robust to using the within-sector employment share instead.

Using this measures, manufacturing industries ranking highly are computer related (SIC 3573, 3571 and 3572), communications equipment (SIC 3669, 3663 and 3661) and electronics and semiconductors (SIC 3672, 3674 and 3676). Since these measures are based on US data only, I will assume the US ranking in the knowledge intensity of industries proxies that of the rest of the world.

Online Appendix Section C provides a discussion on these measures and compares them to other commonly used measures in the literature that proxy for knowledge intensity at the firm level, such as R&D share of expenditures. The O*NET based measures overcome important shortcomings associated with other prevalent measures used in the literature. Their distribution is smoother, they do not rely on a sampling of firms, and they use the same standardized inputs for all industries. Hence, I use these indicators as the main proxies for knowledge intensity throughout the paper.

2.2.3 Unit shipping costs

Unit shipping costs for SIC manufacturing industries are computed using data from Bernard et al. (2006).[14] This industry-level measure aims to proxy the unit shipping cost variable (referred to as t in the theoretical framework in Online Appendix Section A), which accounts for how costly it is to transport one unit of that good irrespective of industry. For instance, goods with the

[14] Downloadable from http://faculty.som.yale.edu/peterschott/files/research/data/xm_sic87_72_105_20120424.zip

highest unit shipping costs in the dataset include ready-mixed concrete and ice, which require special forms of transportation.

The variable measures the amount of US dollars required to transport 1$ worth of a good per every 100Km. It is computed by averaging the same measure per industry across all countries exporting to the US in year 2005. To deal with long tails, this variable will be used in a logarithmic scale in all the different empirical specifications.

Unit shipping cost figures correlate negatively with the knowledge intensity measures, with a correlation coefficient of about -0.6.

2.2.4 Country levels controls

In some specifications I use country level controls such as income per capita from the World Bank's World Development Indicators. In addition, I also often include data on conventionally measured factors of production (stock of physical capital, human capital and land) from UNCTAD (Shirotori et al., 2010).

2.2.5 Ease of communication proxies

In order to proxy for the ease of communication between a subsidiary and its headquarters, I use two variables: non-stop flights and working hours overlap.

The first variable is used because the existence of non-stop flights would proxy for the ease of managers and workers to do more frequent business trips, given the convenience of a direct flight. Business trips, by allowing face-to-face interaction, would facilitate the transmission of tacit knowledge. However, it is important to note that business trips, even if convenient, happen much less often than phone calls due to the elevated costs associated with them. In order to compute the existence of a non-stop air route between a headquarters and its subsidiary, I matched all the existing airports within a 100Km radius (conditional on being in the same country), using the geocoded latitude and longitude. The data for airports (with their respective coordinates) and active air routes come from the OAG Flights[15] Through this matching I create a dummy variable which takes the value of 1 if there is a non-stop flight between the headquarters and its subsidiary.

The second variable, overlap in working hours, aims to capture the "real-time" communication ability between managers and workers in the two plants. Being in the same time zone allows workers to use phone or videoconference communication more frequently (substituting partially for means such as fax or email). This allows for better transmission of tacit knowledge, which is valuable for troubleshooting or crisis solving. In order to compute the overlap in working hours I use the geocoded longitude of each subsidiary to find its time zone, and compare it to that of its headquarters. Assuming that working hours run from 8:00am to 6:00pm (10 hours in total), the variable measures, for a single day, the number of hours that overlap in the working schedule of both the headquarters and its subsidiary.

[15] The dataset was privately acquired in January 2015.

3 Descriptive evidence

In this section I use the global character of the dataset and the "tacit" knowledge intensity measures to estimate two results already established (either empirically or theoretically) in the literature based on the existence of marginal costs to transfer knowledge within the MNC (e.g., costs of monitoring, communicating, traveling, etc.). First, that a MNC is less likely to horizontally expand into a foreign destination the more knowledge intensive is the economic activity under consideration (as shown by Keller and Yeaple (2013) using American MNCs). Second, that conditional on horizontal expansion, there is a trade-off between the knowledge intensity level of the foreign subsidiaries' economic activity and their geographical distance to the headquarters, consistently with what recent frameworks show (e.g., Ramondo and Rodriguez-Clare, 2013; Arkolakis et al., 2013; Keller and Yeaple, 2013; Ramondo, 2014; Tintelnot, 2014).

Table 1 provides some descriptive details about the used sample, in terms of the distribution of foreign affiliates across regions of the world and developing vs. developed countries. This sample includes domestic subsidiaries and foreign subsidiaries that replicate production abroad (i.e. an horizontal expansion). In particular, a foreign subsidiary is included when it reports being active in an industry (a SIC 4 digit code) for which the MNC has one domestic subsidiary also active in.[16] In total, the sample includes 84,881 subsidiaries that are owed by 2030 MNCs. About 32% of these subsidiaries are foreign. The table also includes the knowledge intensity variable measured in standard deviations from the mean (denoted by KI), averaged over domestic and over foreign subsidiaries. The last column presents the difference, with stars denoting the correspondent p-value level.

[Table 1 about here.]

On average, industries of the foreign subsidiaries are less knowledge-intensive than the industries manufactured by their domestic counterparts. The same pattern holds for all presented cuts of the data, besides for few firms based on non-OECD countries, though the difference is not statistically significant. In Western Europe, however, it seems like the pattern is inverted. Beyond the purely descriptive statistics shown in 1 I use plant-level data to estimate the following specification, in an effort to rule out possible confounding factors:

$$Foreign_{h,s} = \beta_k \cdot log(k_s) + \beta_t \cdot log(t_s) + controls_s + \varphi_h + e_{h,s} \qquad (1)$$

Where s indexes a subsidiary and h its headquarters. The independent variable is a dummy which takes the value 1 if the subsidiary is a foreign horizontal affiliate of the firm and 0 if it is a domestic one. k_s is a measure of knowledge intensity of the economic activity (i.e., the manufactured good or product) of the foreign subsidiary. t_s is the unit shipping cost for the good manufactured in the foreign subsidiary. $controls_s$ is a vector of variables that control for the

[16] See Appendix Section B.3 for further explanation on this definition

size of the market and factor endowments of the country hosting the foreign subsidiary which aims to controls for aggregate demand and cost of producing (understanding the limitation of this approach makes this exercise suggestive, rather than definitive).[17] I also include terms to control for whether such foreign horizontal subsidiary also classifies as vertical, either through a downstream or upstream linkage to one of the industries produced domestically by the MNC, allowing for a different constant term for those particular cases.[18] $e_{h,s}$ is the error term. In particular, the idea of this test is to find out whether MNC are less likely to expand internationally their knowledge intensive activities, as a proxy to for worse performance relative to exports (with the proper caveats discussed above).

Using plant-level data provides an additional degree of identification. That is, by adding φ_h, which represents MNC fixed effects, we can control, for example, for the overall productivity level of the MNC. Thus, the results are within-MNC, implying that the partial correlation estimated in β_k is independent of the aggregate productivity level of the MNCs, but only on the economic activity the affiliate is engaged in. It also implies that the comparison is between different subsidiaries active in more or less knowledge intensity activities, within the same MNC. At this point it is worth mentioning that subsidiaries within a single MNC might differ in their economic activity, thus allowing for within-firm variation in the right hand side variables of the empirical specification (see Figure B4 in the Appendix).

Some specifications will also include the term η_s which represents a host country fixed effect (i.e. the country where the affiliate is located);[19] and in separate specifications, the term $\eta_s \times \theta_P$ which represents a fixed effect for host country and 2-digit SIC code (denoted by P) combination. The $\eta_s \times \theta_P$ set of fixed effects will be included to control for issues such as intellectual property regulation, aggregate costs and aggregate demand to the extent they vary at the 2-digit industry level. When both φ_h and η_s are used in the estimation, I exclude the controls defined at the level of the countries of both the headquarters and foreign subsidiary, given multicollinearity.

Before presenting the results it is important to discuss the relevance and interpretation of this particular estimation. In a sense, this estimation aims to estimate whether the level of knowledge intensity of an economic activity plays a role in the proximity-concentration hypothesis. In other words, if knowledge intensive activities are less likely to be expanded horizontally, then this is suggestive evidence that costs associated with the transmission of knowledge are an

[17] These variables include measures of factor endowments between in the subsidiary countries as well as income per capita.

[18] As explained in Section 2.2.1, about 32% of those subsidiaries classified as horizontal would also clarify as vertical according to the Alfaro and Charlton (2009) "5-cents" rule in the input-output relationship. See Appendix Section B.3 for further explanation on this definition. By adding these variables as dummies, we allow for a different constant for these particular subsidiaries.

[19] I slightly abuse from notation when using η_s, which refers to a fixed effect for the host country of the subsidiary, and not for the subsidiary s. I believe this is less confusing than using a new index for the country of the subsidiary.

element driving such trade-off, consistently with others studies –some of them mentioned above– which find a similar result in other contexts. Yet, we also must acknowledge the limitations of this estimation: first and foremost, the absence of exports by each MNC-product combination to particular destinations. Without this, it is very difficult to say something precise about the proximity-concentration trade-off, because it is impossible to know whether the lack of horizontal expansion is happening in a place where the same product is being exported to. Thus, given the limitations, I defer to this descriptive exercise which aims to shed light on whether there is a differential pattern in terms of horizontal foreign expansion with respect to knowledge intensity.[20] The results are presented in Table 2. The table uses the *experience plus training* measure discussed above as a proxy for k.

[Table 2 about here.]

Column 1 presents the preferred specification, which controls for trade unit costs and all the country pair variables. The results suggest that, everything else equal, industries for which their knowledge intensity is 10% higher are about 1.6 percentage points less likely to be replicated abroad. For instance, according to this estimation, semiconductors (SIC 3674), which is characterized by having workers with an average of over 80 months of required experience plus training, is about 12 percentage points -or 21%- less likely to be replicated abroad than a meat packing plant (SIC 2011), which its workers have, on average, 37 months of experience plus training.

The endowment controls suggest that international expansion is more likely in countries that are richer but smaller. It also suggests that it is less likely in countries that are capital abundant and that have high levels of human capital, two variables that typically correlate with higher wages. The estimates for upstream and downstream linkage industries are positive and significant, though besides allowing for a different constant, their interpretation is not straightforward. Finally, it is important to mention that the point estimate for the trade cost variable is negative, though not statistically significant. However, we would have expected a positive estimate according to the proximity-concentration hypothesis. Yet, it is important to notice that this unexpected result is probably due to the inclusion of multinational fixed effects, given that most of the variation of the trade cost variable is across MNCs and not within. Thus, the inclusion of MNCs fixed effects comes at the price of losing the ability to exploit the variation of this variable. Alternative standardizations of the trade cost variable often provide positive point estimates, though still statistically insignificant. I discuss this particular result in detail later on.

Consistent with the idea that the transmission of knowledge plays a role in the proximity-concentration hypothesis, the estimator for β_k is negative and robust across all specifications. The inclusion of host country fixed effects denoted

[20] An alternative approach would be to "fill" with zeros for each country-product combination where there is no horizontal expansion, but this implies assuming that there is (equal) demand everywhere, which is a very strong assumption.

by η_s in Column 3 rules out other potential channels that could be driving the results. For instance, poor intellectual property regulation in different countries.[21] Column 4 includes a more strict control: parent industry (2-digit), denoted by P, interacted with host country fixed effects, $\eta_s \times \theta_P$, to allow for differential policies at the country level for different types of industries. Note that in columns 2 and 3 we exclude the country level controls due to high multicollinearity. Overall, the estimate of β_k is robust to the inclusion of these additional controls in terms of its magnitude, negative sign and its statistical significance.[22]

The second implication of the existence of marginal costs of transferring knowledge within the MNC are that foreign affiliates locate in closer geographic proximity to the headquarters the more intensive in knowledge their main economic activity is (assuming these costs are increasing in distance). To answer this question I restrict the dataset to only existing *foreign* subsidiaries, leaving about 25 thousand subsidiaries in the sample (i.e. I exclude observations which correspond to a domestic subsidiary of the MNC). Table 3 presents summary statistics for this foreign subsidiaries, including the geographic distance in between the headquarters and the subsidiary itself, the knowledge intensive and unit shipping cost values as well as other control variables. For instance, the average distance between a headquarters and its foreign subsidiary in the sample is about 8.2 in logs (in levels, it corresponds to about 5600 Km). The table also presents the average knowledge intensity for economic activity of the foreign subsidiary, as well as the corresponding unit shipping cost (which distributes between 0 and 1, and thus its logarithmic transformation is always negative). As explained above, some of the foreign subsidiaries that classify as horizontal, *also* classify as vertical (both downstream and/or upstream), and thus in all the specifications I allow for a different constant term by adding dummies for these foreign subsidiaries. In particular 29% (15.3%) of the horizontal foreign subsidiaries are in an economic activity that could be upstream (downstream) to the economic activities of the domestic subsidiaries. There is an overlap between the two, so in total, about 30% of the foreign subsidiaries would classify as horizontal and vertical (either upstream or downstream). The lower panel include summary statistics for variables that measure the ease of communication between the headquarters and the foreign subsidiary, which will be used in the next subsection. For instance, about 33% of the headquarters and foreign subsidiary pairs have a non-stop route between their nearby airports.[23] The table also shows that, on average, there are almost 7 working overlapping hours between a headquarters and its subsidiary.

[21] Results are robust to excluding China from the sample, alleviatng possible concerns given biases its inclusion might generate in the estimations due lack of enforcement with respect to intellectual property rights.

[22] This exercise is similar to the one by Keller and Yeaple (2013), though not quite, as it does not control for local demand in the foreign location. In Online Appendix Section D, however, I estimate the same relationship as in Keller and Yeaple (2013) using aggregated data on exports and sales of foreign affiliates of US companies by industry, making use of the knowledge intensity measures I constructed.

[23] Using 2012 data, the same year the Dun and Bradstreet data is representative of.

[Table 3 about here.]

Figure 3 show graphic results suggesting the existence of a trade-off between distance to the headquarters and knowledge intensity of the economic activity of the foreign subsidiary. It plots the estimation of β_k using the following specification:

$$log(d_{h,s}) = \beta_k \cdot log(k_s) + \beta_t \cdot log(t_s) + \varphi_h + e_{h,s} \tag{2}$$

Where s indexes a subsidiary and h its headquarters. The left hand side variable, $log(d_{h,s})$ is the logarithmic transformation of the geographic distance between the location of the headquarters and that of the foreign subsidiary. The right hand side includes the log of knowledge intensity associated to the economic activity of the subsidiary (denoted by k_s), the log unit shipping cost (t_s), as well MNC fixed effects. We would expect $\beta_k < 0$: foreign subsidiaries will tend to be closer the more knowledge intensive they are (given that the transmission of knowledge is increasing in distance). Note that the inclusion of MNC fixed effects, imply that the exploited variation *within* firm.

The results are indicative of the trade-off between distance (d) and knowledge intensity (k). In fact, according to the estimation, for every 10% increase in the knowledge intensity scale, a foreign subsidiary is about 5% closer in distance to the headquarters.[24] In practice, the results suggest that, roughly, an American MNC headquartered in the Houston, Texas would locate its low knowledge intensive activity, say a meat packing subsidiary, in Kabul, Afghanistan (approximately, 12000Km), while a high knowledge intensive activity, for example, a semiconductor plant, would be located in a place that is about 40% closer, so, for instance, in Ireland (approximately 7000km distance), *ceteris paribus*.

[Figure 3 about here.]

To avoid any misunderstanding coming from using the term "knowledge and distance trade-off", an explanation is in place. The term refers to the finding that –conditional on foreign expansion– MNCs tend to locate their foreign subsidiaries active in knowledge intensive activities closer to the headquarters than those active in less knowledge intensive activities. This result is an extension of the traditional proximity-concentration hypothesis in the presence of marginal costs of transferring knowledge within the MNC that increase in distance. Thus, in a way, the result is could be interpreted as part of the proximity-concentration hypothesis.

[24] Allowing for a more flexible fit, such as a quadratic one, suggests that the estimated quadratic relationship does not seem to be monotonically decreasing for the lower values of knowledge intensity (although a flat or even negative slope in that area cannot be rejected in the data either). However, and perhaps more importantly, for higher levels of knowledge intensity there is a clear negative relationship with distance. This result is qualitatively important, given that it would be consistent with the idea that distance appears to matter much more for higher levels of knowledge intensity. Intuitively, this means that after certain level of knowledge intensity, the more sophisticated products are the closer the foreign subsidiaries will be located to the headquarters. Results of this exercise are available upon request.

4 The distance-knowledge trade-off and the ease of communication

With this evidence in hand, it might be difficult to attribute the documented trade-off between k and d to the somewhat abstract idea of marginal costs of transferring knowledge. The empirical results might be driven by factors other than knowledge not accounted for, in the presence of omitted variable bias. A conventional explanation in the literature would be that knowledge intensive sectors are associated with higher intra-firm trade of intermediate goods, making it less profitable to locate those plants in far away locations (Keller and Yeaple, 2013; Irarrazabal et al., 2013). Keller and Yeaple (2013), in particular, assume in their model that knowledge can be fully embedded in intermediate goods, that are in turn shipped to remote locations.

However, this assumption might not be proper for all type of knowledge. Tacit knowledge cannot be, by definition, embedded in intermediate goods. For instance, transmission of tacit knowledge within the firm could involve costs of monitoring by the management in the headquarters or also real-time problem solving efforts by the chief engineering team dealing with the foreign team in certain manufacturing processes. Distance would play an important role in this matter: business travel, for instance, or the ability to work in teams in real time (through phone, video conferences, etc.) would be critical for industries where monitoring and work across borders is prevalent. Thus, it could well be that it is the cost of transmitting this type of knowledge which partly drives the documented relationship, and not the shipping of intermediate goods.

I perform a test that aims to disentangle between both explanations. If the cost of transferring knowledge is indeed an increasing function of distance -as argued- and thus, a determinant in the location decisions of firms, then easier communication between headquarters and subsidiaries would work as a cost-reducing mechanism for the purpose of transmitting knowledge. This would be hard to explain with the intra-firm trade mechanism, given that arguably the ease of communication is orthogonal to the transportation costs of intermediate goods.

I test for this hypothesis by re-estimating specification (2), this time including variables that proxy for the ease of communication within the firm, often interacting them with the knowledge intensity covariate. These variables, all measured for each subsidiary and its headquarters, are (1) the existence of a commercial non-stop air route (between airports close to the headquarters and the foreign subsidiary, measured as being within 100Km and within the country borders), and (2) the number of overlapping working hours in a business day (see Section 2.2.5 above for more details on the construction of these variables).

The purpose of utilizing these variables is to proxy for forms of communication that allow for the transmission of knowledge, though they are quite different between themselves. As explained above, business travel provides the opportunity to work face-to-face, though it occurs with less frequency, given the high costs of traveling. In this context, Giroud (2012) has shown that the existence of commercial air routes between subsidiaries and headquarters positively affects the profitability of the former. Being in the same time zone allows for convenient real-time, day-to-day, communication, significantly reducing waiting time between the two ends for problem solving or consulting about specific tasks. Stein and Daude (2007), for instance, find that time zone is an important determinant of aggregate FDI flows, which they attribute to better monitoring.

The results are presented in Table 4. All columns use the *experience plus training* indicator to proxy for k. Columns 1 and 2 present result using the existence of a non-stop flight between the headquarters and the foreign subsidiary, while Columns 3 and 4 present results using the overlap in working hours between the foreign subsidiary and its global headquarters (the value goes from 0 to 10).

[Table 4 about here.]

First, across all specifications, the estimates for β_k are negative and statistically significant, signaling that the trade-off between distance and knowledge intensity remains there, on average.[25] Column 1 shows that neither the dummy variable measuring the existence of a non-stop flight in between the headquarters and its foreign subsidiary, nor its interaction with the knowledge intensity measure are statistically different from zero. In fact, the estimate of β_k is not very different than the one estimated and plotted in Figure 3. Column 2 repeats the exercise including a fixed effect for the country of the foreign subsidiary, and thus the variables controlling for factor endowments in those countries are excluded. The results, however, are robust to this inclusion, yet they are harder to interpret because of the variation left on the dependent variable after both fixed effects being added simultaneously.[26]

Columns 3 and 4 show a different result when focusing on overlap in working hours as a measure of ease in communication between the headquarters and the foreign subsidiary. In fact, there the estimate for β_k becomes significantly larger maintaining its negative sign, implying that when there are no overlap in working hours, the distance and knowledge intensity trade-off becomes much more pronounced. In addition, and perhaps more importantly, the interaction

[25] All specifications control for the trade unit costs, and the estimator for β_t is not significantly different from zero. While we would have expected a positive point estimate, the results are not surprising as we only exploit the within MNC variation. The other controls imply that distance from headquarters to foreign affiliate correlates positively with both population and human capital levels and negatively with capital per worker levels of the foreign affiliate's host country. It is not straightforward to interpret these numbers beyond their role as control, but one possible interpretation in this context is that higher level of human capital in the affiliate's country correlates with better local management, and thus the cost of transferring knowledge is reduced, allowing for more distant subsidiaries. Larger population and lower levels of capital per workers is correlated with lower wages, too, which might play a role in locating subsidiaries in more distant locations. The estimate for upstream and downstream linkages dummy variables do not statistically differ from zero, and the results are robust to their exclusion.

[26] In part this is the reason I don't report results that include the $\eta_s \times \theta_P$ fixed effects as in previous tables. The results, however, are robust to the inclusion of this additional control.

term appears to be positive and statistically significant, implying that the trade-off weakens the more overlap in working hours. It is also worth noting that across all specifications the ease of communication coefficients are estimated to be negative, and this is simply because the longer the distance between these two countries, these variables tend to be smaller (i.e. less working hours overlap, less non-stop flights).

These findings are insightful on their own. The results suggest that being in the same time zone seems to facilitate the transmission of knowledge. The ability of managers in the headquarters to communicate with colleagues in foreign locations, for troubleshooting or consulting on an open-ended range of issues, is more efficient when communication happens in real time, without long waiting times. This might be even more relevant for transmitting tacit knowledge, given that complicated problems would require real-time interaction, and not just explanations being sent through fax or email. Furthermore, this logic could also serve as an example for arguing that the barriers of transmitting knowledge is increasing with distance: managers and workers in the headquarters might require working extra hours to communicate with their peers in foreign subsidiaries, incurring additional compensation and operational costs.

Figure 4 presents a graphical representation of the distance and knowledge intensity trade-off for different the different possible values that define the overlap in working hours between the foreign subsidiary and its headquarters. The left panel plots the estimated slope of such trade-off as a function of overlap in working hours. When the overlap is 0, the slope of the trade-off is highly negative. As the overlap becomes larger, the trade-off weakens and even disappears. The right panel, plots the trade-off for different values of overlap in working hours (normalizing the $log(k)$ to be zero at its smallest value of the distribution, for comparison purposes), showing that the trade-off is even slightly positive when there are 10 hours of overlap between the two locations. Simply put, this result implies that the cost of shipping intermediate goods, which would be just as relevant within the same time zone (because north-south shipping is equally as expensive as east-west shipping) is not enough to explain the fact that MNCs tend to locate their foreign subsidiaries active in knowledge intensive geographically nearby.

[Figure 4 about here.]

A regression discontinuity design to study the role of time zones

In order to rule out biases arising from confounding factors, I implement a regression discontinuity framework exploiting the fact that time zones vary discretely with distance. By doing so, I effectively compare the knowledge intensity of foreign subsidiaries that belong to the same MNC but that are on different sides of a time zone line.[27] To the best of my knowledge, using a regression

[27]I refrain from presenting results using a fuzzy regression discontinuity design using the existence of a non-stop flight based on the work by Campante and Yanagizawa-Drott (2016),

discontinuity design exploiting time zone differences is novel in the context of international economics, and in particular, when studying patterns of behavior of multinational corporations.

As a first step it is important to define the running variable centered around each time zone line: it corresponds to the distance (in kilometers) from every foreign subsidiary to the nearest time zone line. A MNC will often have foreign subsidiaries in located in time zones that are both eastwards and westwards relative to the headquarters location. Our "treatment" in this setting is being in a time zone that allows for a higher overlap in working hours between the headquarters and the subsidiary. Figure 5 clarifies how the running variable is defined and normalized using an hypothetical example. For instance, foreign subsidiaries 1 and 2 are located east of their headquarters' location, while subsidiaries 3 and 4 are located west of it. The running variable is based on the distance from each one of these subsidiaries to the nearest time zone line. In that sense, because foreign subsidiary 1 is located on the side of the line closest to the headquarters (e.g., higher overlap in working hours), its distance is marked as positive. On the other hand, since foreign subsidiary 2 is located on other side of the line, then the running variable in that area is negative. Same concept applies to the running variable for foreign subsidiaries 3 and 4.

[Figure 5 about here.]

Each observation in the sample corresponds to a foreign subsidiary, and thus the "treatment" can be defined as being in a closer time zone to its headquarters. Each headquarters, of course, has several foreign subsidiaries that are near several distinct time zone lines. Thus, in this setting it is very important to make sure that the regression discontinuity design it is indeed comparing foreign subsidiaries that belong to the same MNC and at the same time, that are near the same time zone line to one side or the other. Thus, I estimate the following specification:

$$\begin{aligned} log(k_{h,s}) = \beta_{TZ} closerTZ_{h,s,tz} + distTZline_{h,s,tz} + \\ closerTZ_{h,s,tz} \times distTZline_{h,s,tz} + \varphi_h + \tau_{tz} + e_{h,s} \end{aligned} \quad (3)$$

Where h is a subscript for each MNC, s represents a subsidiary and tz represents a time zone line which is the nearest one to subsidiary s. $log(k_{h,s})$ is the knowledge intensity of the economic activity of the foreign subsidiary. The variable is a dummy which is defined as $closerTZ_{h,s,tz} = 1\{distTZline_{h,s,tz} >$

who show that the likelihood of having a non-stop flight is experiences a jump at the 6000 miles (or 9600 kilometers) of distance. The reason not to present those results are twofold. First, the exercise shown in Table 4 finds no results when including a non-stop flight as an ease of communication variable. Second, while the reduced form (using the 6000 mile threshold) shows results consistent with what expected (e.g., observations just below the threshold are more knowledge intensive than those just above it), the results are not robust when using the non-stop flight in a fuzzy regression discontinuity setting. This is probably because the sample does not include all the existing flights, but rather only those flying to and from airports near a headquarters or foreign subsidiaries. Results of this, however, are available upon request.

0}. $distTZline_{h,s,tz}$ is the running variable, defined as distance from subsidiary s to its closest time zone line tz following the guidelines explained before and represented in Figure 5. The interaction $closerTZ_{h,s,tz} \times distTZline_{h,s,tz}$ allows more flexibility when estimating the slopes before and after the cutoff (e.g., the time zone line). φ_h and τ_{tz} are MNC and time zone line fixed effects, respectively. By adding these fixed effects I make sure that we compare foreign subsidiaries belonging to the same MNC and that are across the same time zone line.

Figure 6 graphically represents the estimation of Specification (3) using the package *rdplot* (Calonico et al., 2014, 2017). It shows that, indeed, that foreign subsidiaries located just across the time zone line that puts them in a closer time zone to their headquarters (above zero in the horizontal axis which measures values of the running variable) are, on average, are active in economic activities higher in the knowledge intensity scale (measured in the vertical axis), as compared to those who are on the other side of the time zone (below zero in the horizontal axis). This graph uses observations for which their distance to the time zone line corresponds to the optimal running variable bandwidth which is 299.157 kilometers, using the methodology by Cattaneo et al. (2018) who builds on the work by Imbens and Kalyanaraman (2012). Note that the vertical axis has values below and above zero because for this plot I use the residual of $log(k_{h,s})$ after controlling for φ_h and τ_{tz}.

[Figure 6 about here.]

The actual estimation of Specification (3) using data on foreign subsidiaries are presented in Table 5. Each column in the table presents results using a different bandwidth definition which range from 150 to 350 kilometers as well as the optimal bandwidth in the last column. The estimation uses a triangular weight scheme, giving higher weight to observations closer to the cutoff point.

[Table 5 about here.]

The results suggest that a foreign subsidiary just across (any) time zone line closer to the headquarters –thus increasing the overlap in working hours– is in an economic activity that is, on average, 0.6 to 0.84 percent higher in the knowledge intensity scale, depending on the bandwidth used. Note that the estimator for β_{TZ} is statistically significant (the standard errors are clustered at the MNC and time zone line level). At first this might look like a very small number. Yet, keep in mind that the treatment itself is also quite small: this exercise compares foreign subsidiaries very close to each other (certainly measured in terms relative of distance to their headquarters), some of them (the treated) have an extra hour of overlap with the work day of their colleagues at the headquarters, or sometimes even less than that (some time zones change in increments of half an hour). Also, this is an average treatment effects for all possible different treatments (e.g., increasing the overlap in working hours from 2 to 3, or from 6 to 7, for example) and as such is a weighted average of different treatment effects which could be larger or smaller, depending on the relative improvement.

Yet, at most we will always be limited by a "small" treatment. Thus, the results are qualitatively important taking into consideration that the only thing which differs between subsidiaries located at either side of the time zone line is having one more hour of overlap with their colleagues working at the headquarters location, even if their external validity cannot be proven beyond the estimation of presented in Table 4, with all the limitations such approach has.

However, a question remains: is it overlap in working hours the only thing that varies substantially when comparing subsidiaries at either time of a time zone line? The evidence suggest that, in the context of what is relevant for this exercise, it is. To show that I re-estimated Specification (3) replacing the dependent variable for other observable measures considered determinants of the location decision of foreign subsidiaries (used as controls in several specifications above). These are the trade unit cost and measures of factor endowments in the country where the subsidiary is located (income per capita, capital per worker and human capital). Since some of the horizontal foreign subsidiaries in the sample could also be classified as being upstream or downstream to the portfolio of economic activities produced at home by the MNC, I also test for those. The results can be found in Table 6 where each column uses a different dependent variable in the context of the regression discontinuity (using the optimal bandwidth and a triangular weighting scheme). As can be seen, the treatment effect is statistically not different from zero across all of those variables, reducing remaining concerns of other confounding factors: time zone changes cannot explain differential patterns in terms of trade costs (Column 1), nor whether the horizontal foreign subsidiary is being misclassified (Columns 2 and 3), and neither in the factor endowments of the country where the subsidiary is located (Columns 4, 5 and 6).

[Table 6 about here.]

5 Concluding Remarks

It is straightforward that, in order to operate smoothly, MNCs make investments so that their workers located both in the headquarters and in foreign subsidiaries can work together to maintain a smooth business operation. "Working together" could imply many things, such as managing, coordinating, monitoring, troubleshooting, etc. One broad way to denominate all these activities is under the umbrella of knowledge transmission. Knowledge transmission can certainly be costly and more intense for certain industries. These costs, of course, are an important aspect determining location decisions of MNCs. This reasoning is, of course, the main explanation behind the existing trade-off between distance to the headquarters and the knowledge intensity level of a foreign subsidiary, a result consistent with the work of many others (e.g., Ramondo and Rodriguez-Clare, 2013; Ramondo, 2014; Arkolakis et al., 2013; Tintelnot, 2014; Keller and Yeaple, 2013).

The main contribution of this paper, however, is to show that this trade-off significantly weakens when time zones are taking into consideration: knowledge intensity is, on average, higher for roughly equidistant foreign subsidiaries that are closer to the headquarters in terms of in time zones. This result is at odds with suggesting that larger intra-firm trade costs (between the headquarters and its knowledge intensive subsidiaries) can fully explain the aforementioned trade-off (e.g., Keller and Yeaple, 2013). Some other distance-varying costs must be playing a role in explaining this trade-off. In particular costs that can be reduced by engaging in real time communication between different locations. A possible interpretation is the intra firm transmission of tacit knowledge.

This paper also contributes to the literature by adapting a regression discontinuity design that exploits multiple spatial differences across time zone lines in the context of researching the behavior of MNCs. To the best of my knowledge, this is the first time this methodology has been used in this context.

Nonetheless, this paper has left open some other specific questions that will shed light on our general understanding of how the transmission of knowledge plays a role in the location decisions and the performance of MNC. For instance, what other tools and means are at a firm's disposal to enhance the process through which it transfers knowledge to its subsidiaries and workers?[28] How does the knowledge intensity of the good relate to the existence of regional hubs, as opposed to different plants serving every foreign market? Is the cost of knowledge transmission a relevant determinant for service provider firms, as it is for manufacturing firms? Given the difference in the nature of services vs. manufacturing industries in terms of their tradability, we can expect different patterns in the data.

Naturally, this research agenda also contains questions that have relevant policy implications. While governments intend to develop their private sectors by attracting foreign investment, designing an effective policy should answer questions such as: is there enough infrastructure in place to allow effective communication for foreign firms? Should the focus be on specific types of firms and specific industries for which knowledge transmission will be easier? Do all types of products have the potential to generate productivity spillovers to domestic firms, or only those for which the cost of knowledge transmission is low?

All in all, despite the fact that productivity outweighs factor accumulation in growth accounting exercises (e.g., Hall and Jones, 1999; Caselli, 2005), the process through which knowledge is accumulated by economic agents is still an under-researched area. However, a better understanding of this process is critical to answering open questions in economics. The difficulties associated with transferring and acquiring knowledge, which translates into productivity shifts, are not unique to MNCs. They can also relate to domestic firms (e.g., Bloom et al., 2012; Kalnins and Lafontaine, 2013), investors (e.g., Coval and Moskowitz, 2001), innovation (e.g., Kerr, 2008) and even countries' export baskets diversification (e.g., Bahar et al., 2014; Bahar and Rapoport, 2018). At a larger scale, the documented evidence reinforces the importance of knowledge

[28]Using German data Gumpert (2015) shows that MNCs active in knowledge-intensive activities "distribute the knowledge" by hiring more capable workers to foreign subsidiaries.

transmission in overall economic activity. Thus, understanding the ways knowledge affects economic activity lies at the core of important and unanswered questions on convergence, development and growth. Knowledge and its diffusion, after all, are significant phenomena that can alter global economic patterns in as-of-yet unexplored ways.

References

Acemoglu, Daron, Simon Johnson, and Todd Mitton. “Determinants of Vertical Integration : Financial Development and Contracting Costs.” *The Journal of Finance* LXIV, 3: (2009) 1251–1290.

Alfaro, Laura, Pol Antras, Davin Chor, and Paola Conconi. “Internalizing Global Value Chains: A Firm-level Analysis.” *NBER Working paper* 41, December: (2015) 1–5.

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

Alfaro, Laura, and Maggie Xiaoyang Chen. “Surviving the Global Financial Crisis: Foreign Ownership and Establishment Performance.” *American Economic Journal: Economic Policy* 4, 3: (2012) 30–55.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. “Innovation and Production in the Global Economy.” Technical Report 18792, National Bureau of Economic Research, Cambridge, MA, 2013.

Arrow, Kenneth J. “Classificatory Notes on the Production and Transmission of Technologcal Knowledge.” *The American Economic Review* 59, 2: (1969) 29–35.

Autor, David H, Frank Levy, and Richard J Murnane. “The Skill Content of Recent Technological Change: An Empirical Exploration.” *The Quarterly Journal of Economics* 118, 4: (2003) 1279–1333.

Bahar, Dany. “Measuring knowledge intensity in manufacturing industries: a new approach.” *Applied Economics Letters* .

Bahar, Dany, Ricardo Hausmann, and Cesar A. Hidalgo. “Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?” *Journal of International Economics* 92, 1: (2014) 111–123.

Bahar, Dany, and H. Rapoport. “Migration, Knowledge Diffusion and the Comparative Advantage of Nations.” *The Economic Journal* .

Becker, Randy A, Wayne B Gray, and Jordan Marvakov. “NBER-CES Manufacturing Industry Database: Technical Notes.” Technical Report February, 2013.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. "Trade costs, firms and productivity." *Journal of Monetary Economics* 53, 5: (2006) 917–937.

Black, SE, and PE Strahan. "Entrepreneurship and bank credit availability." *The Journal of Finance* LVII, 6: (2002) 2807–2833.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics* 128, 1: (2012) 1–51.

Bottazzi, Laura, and Giovanni Peri. "Innovation and spillovers in regions : Evidence from European patent data." *European Economic Review* 47, 4: (2003) 687–710.

Brainard, SL. "A simple theory of multinational corporations and trade with trade-off between proximity and concentration." *NBER Working Paper Series* , 4269.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. "Rdrobust: Software for regression-discontinuity designs." *Stata Journal* 17, 2: (2017) 372–404.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. "Robust data-driven inference in the regression-discontinuity design." *Stata Journal* 14, 4: (2014) 909–946.

Campante, Filipe, and David Yanagizawa-Drott. "Long-Range Growth: Economic Development in the Global Network of Air Links." *NBER Working Paper* No. 22653: (2016) 50.

Carr, DL, JR Markusen, and KE Maskus. "Estimating the knowledge-capital model of the multinational enterprise." *The American Economic Review* 91, 3: (2001) 693–708.

Caselli, Francesco. "Accounting for cross-country income differences." *Handbook of Economic Growth* 1, 05: (2005) 679–741.

Cattaneo, Matias D, Nicolas Idrobo, and Rocío Titiunik. "A Practical Introduction to Regression Discontinuity Designs." *Cambridge Elements: Quantitative and Computational Methods for Social Science - Cambridge University Press* I.

Caves, RE. "International Corporations: The Industrial Economics of Foreign Investment." *Economica* 38, 149: (1971) 1–27.

Costinot, Arnaud, Lindsay Oldenski, and James Rauch. "Adaptation and the boundary of multinational firms." *The Review of Economics and Statistics* 93, February: (2011) 298–308.

Coval, Joshua D., and Tobias J. Moskowitz. “The Geography of Investment: Informed Trading and Asset Prices.” *Journal of Political Economy* 109, 4: (2001) 811–841.

Fan, JPH, and LHP Lang. “The Measurement of Relatedness : An Application to Corporate Diversification.” *The Journal of Business* 73, 4: (2000) 629–660.

Giroud, Xavier. “Proximity and Investment: Evidence from Plant-Level Data.” *The Quarterly Journal of Economics* 128, 2: (2012) 861–915.

Gumpert, Anna. “The Organization of Knowledge in Multinational Firms.” http://papers.ssrn.com/abstract=2622415.

Hall, R. E., and C. I. Jones. “Why do Some Countries Produce So Much More Output Per Worker than Others?” *The Quarterly Journal of Economics* 114, 1: (1999) 83–116.

Harrison, Ann E., Inessa Love, and Margaret S. McMillan. “Global capital flows and financing constraints.” *Journal of Development Economics* 75, 1: (2004) 269–301.

Helpman, E. “A Simple Theory of International Trade with Multinational Corporations.” *The Journal of Political Economy* 92, 3: (1984) 451–471.

Helpman, Elhanan, Marc J Melitz, and Stephen R Yeaple. “Export Versus FDI with Heterogeneous Firms.” *American Economic Review* 94, 1: (2004) 300–316.

Imbens, Gguido, and Karthik Kalyanaraman. “Optimal bandwidth choice for the regression discontinuity estimator.” *Review of Economic Studies* 79, 3: (2012) 933–959.

Irarrazabal, Alfonso, Andreas Moxnes, and LD Opromolla. “The Margins of Multinational Production and the Role of Intrafirm Trade.” *Journal of Political Economy* 121, 1: (2013) 74–126.

Jaffe, A.B., M. Trajtenberg, and R. Henderson. “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations.” *The Quarterly Journal of Economics* 108, 3: (1993) 577.

Kalnins, Arturs, and Francine Lafontaine. “Too Far Away? The Effect of Distance to Headquarters on Business Establishment Performance.” *American Economic Journal: Microeconomics* 5, 3: (2013) 157–179.

Keller, Wolfgang. “Geographic localization of international technology diffusion.” *American Economic Review* 92, 1: (2002) 120–142.

———. “International Technology Diffusion.” *Journal of Economic Literature* XLII, September: (2004) 752–782.

Keller, Wolfgang, and Stephen Ross Yeaple. "The Gravity of Knowledge." *American Economic Review* 103, 4: (2013) 1414–1444.

Kerr, W R. "Ethnic Scientific Communities and International Technology Diffusion." *Review of Economics and Statistics* 90, 3: (2008) 518–537.

Lipsey, RE. "The Creation of Microdata Sets for Enterprises and Establishments." *Annales de l'INSEE* , 30: (1978) 395–422.

Markusen, James R. "Multinationals, multi-plant economies, and the gains from trade." *Journal of international economics* 16, 1984: (1984) 205–226.

Markusen, James R., and Keith E. Maskus. "Discriminating among alternative theories of the multinational enterprise." *Review of International Economics* 10, 4: (2002) 694–707.

Markusen, JR. "Trade versus Investment Liberalization." *NBER Working Paper Series* , 6231.

Markusen, JR, AJ Venables, DE Konan, and KH Zhang. "A unified treatment of horizontal direct investment, vertical direct investment, and the pattern of trade in goods and services." *NBER Working Paper Series* , 5696.

Nunn, Nathan, and D Trefler. "The boundaries of the multinational firm: an empirical analysis." In *The Organization of Firms in a Global Economy*, edited by E Helpman, D Marin, and T Verdier, Cambridge, MA: Harvard University Press, 2008, 55–83.

Oldenski, Lindsay. "Export Versus FDI and the Communication of Complex Information." *Journal of International Economics* 87, 2: (2012) 312–322.

Pierce, JR, and PK Schott. "A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries." *Journal of Economic and Social Measurement* , 301.

Polanyi, M. *Personal knowledge: Towards a post-critical philosophy.* London, UK: Routledge, 1962.

———. *The Tacit Dimension.* Chicago; London: University of Chicago Press, 1966, 2009 edition.

Pritchett, Lant. "Divergence, big time." *The Journal of Economic Perspectives* 11, 3: (1997) 3–17.

Ramondo, Natalia. "A quantitative approach to multinational production." *Journal of International Economics* 93, 1: (2014) 108–122. `http://dx.doi.org/10.1016/j.jinteco.2014.01.004`.

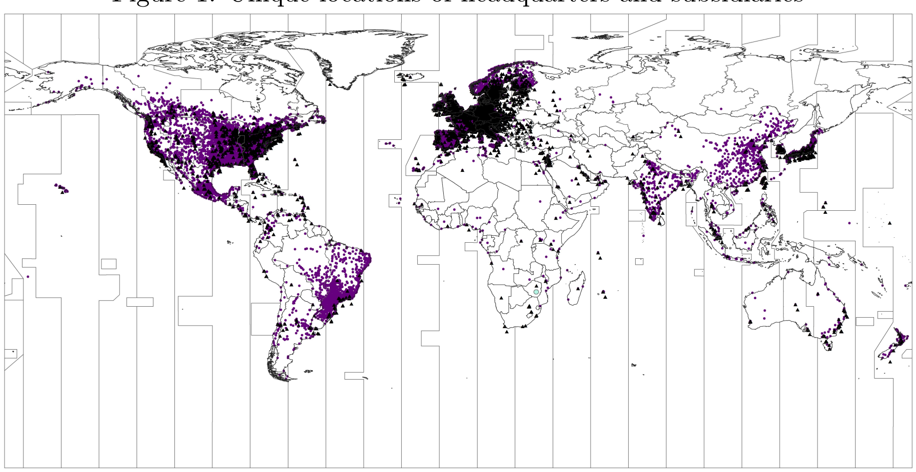
Ramondo, Natalia, and Andrés Rodriguez-Clare. "Trade, Multinational Production, and the Gains from Openness." *Journal of Political Economy* 121, 2: (2013) 273–322.

Shirotori, M, B Tumurchudur, and O Cadot. “Revealed Factor Intensity Indices at the Product Level.” *Policy Issues in International Trade and Commodities* 2010, 44.

Stein, Ernesto, and Christian Daude. “Longitude matters: Time zones and the location of foreign direct investment.” *Journal of International Economics* 71, 1: (2007) 96–112.

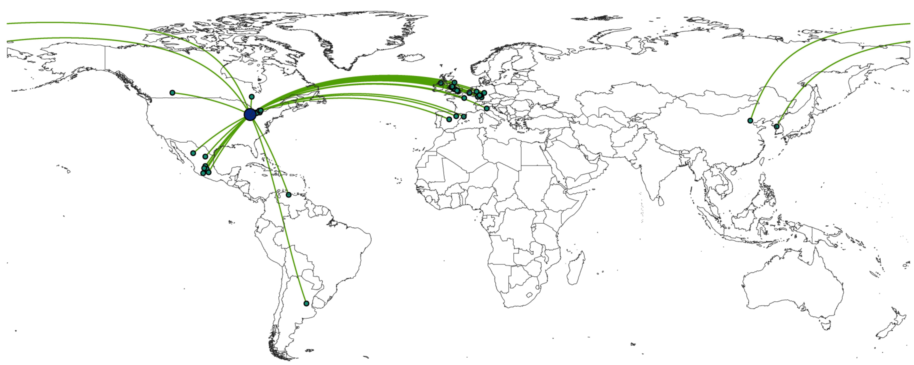
Tintelnot, Felix. “Global Production with Export Platforms.”, 2014.

Figure 1: Unique locations of headquarters and subsidiaries



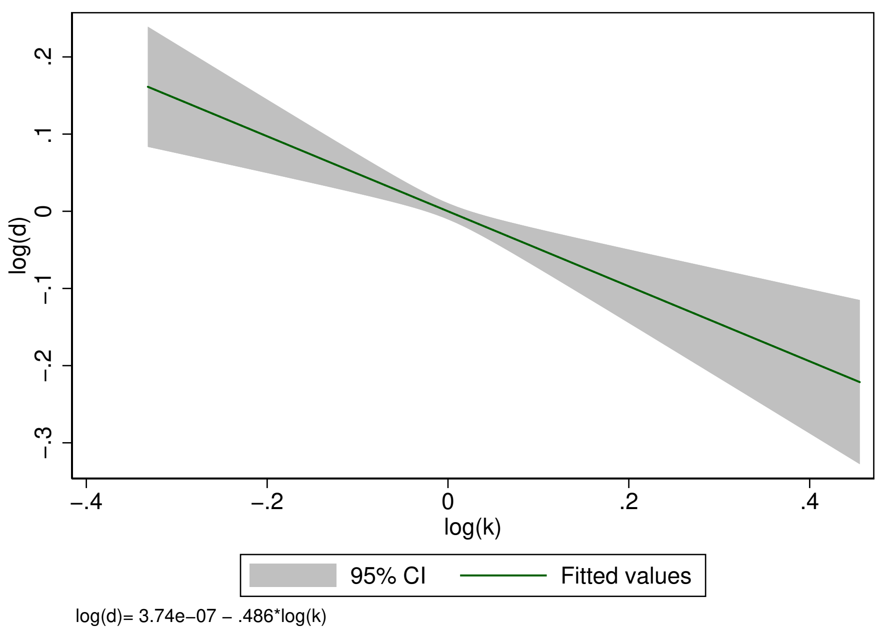
The figure shows a World map with the geocoded location of all the headquarters (triangles) and foreign subsidiaries (dots) in the sample.

Figure 2: Headquarters and foreign subsidiaries of an American MNC



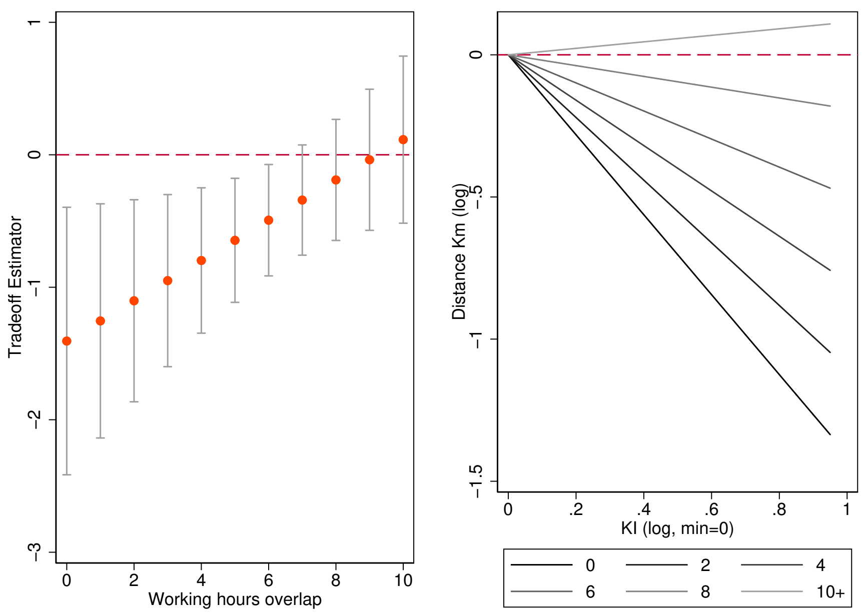
The figure is an example of the resolution of the data. It shows a World map with the geocoded location of the headquarters of an American car manufacturing firm and all of its subsidiaries.

Figure 3: Estimated relationship of k and d



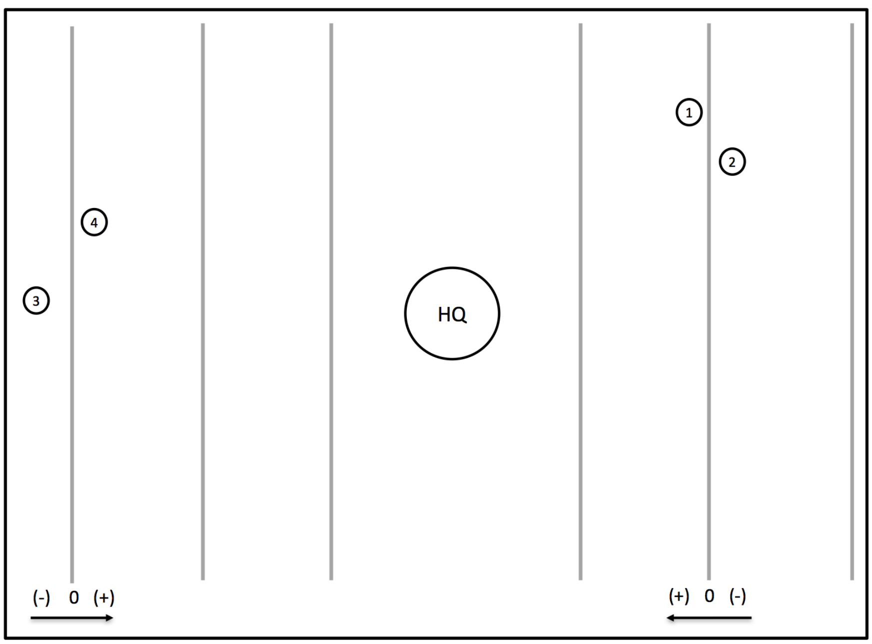
The figure presents the empirical fit for the relationship between $log(d)$ and $log(k)$ (the latter proxied by the *experience plus training* measure), after controlling for unit trade costs and MNC fixed effects. 95% confidence intervals marked in grey, based on robust standard errors clustered at the MNC and industry level.

Figure 4: KI and distance trade-off by working hours overlap



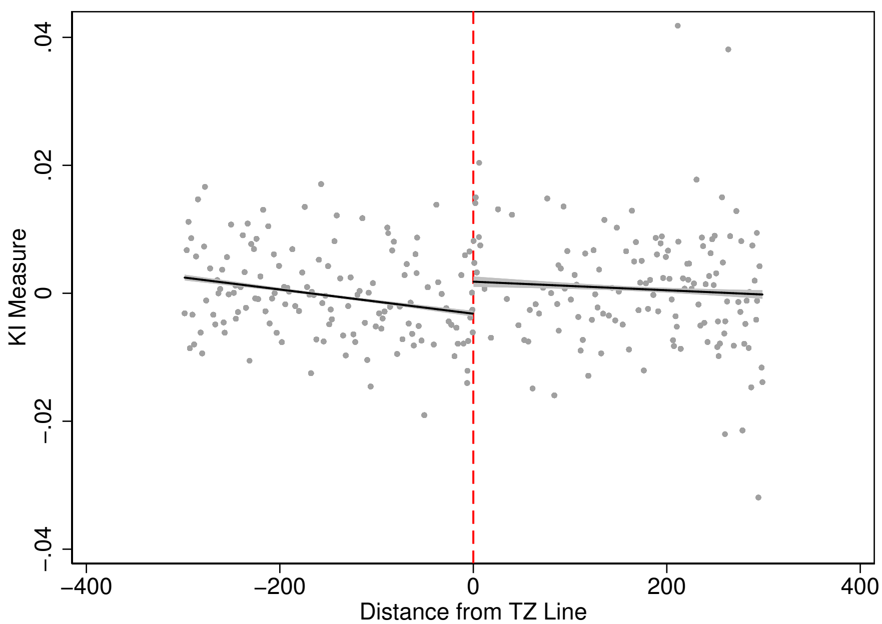
The figure presents the empirical fit for the relationship between $log(d)$ and $log(k)$ (the latter proxied by the *experience plus training* measure), after controls specified in Specification (\eqref{eq:dd-dk}), and adding different controls for the ease of communication between the headquarters and the foreign affiliate.

Figure 5: Running variable definition



The figure presents a graphical representation to understand how the running variable (distance to the nearest time zone line) is centered around the time zone line.

Figure 6: Regression discontinuity graphical representation



The figure presents a graphical representation of the regression discontinuity design in Specification 3.

Table 1: Descriptive Statistics (Domestic Vs. Foreign Subsidiaries)

	MNC	# Subs	Foreign (%)	$KI_{Foreign}$	$KI_{Domestic}$	Δ
All Observations	2030	84881	.32	.27	.34	-.079***
Non OECD	73	2445	.13	.39	.37	.027
OECD	1957	82436	.32	.26	.34	-.079***
East Asia & Pacific	396	20579	.13	.45	.44	.0023
Latin America & Caribbean	30	2247	.56	-.36	-.34	-.017
North America	714	33679	.24	.36	.38	-.016*
South Asia	43	1557	.13	.38	.45	-.067*
Western Europe	847	26819	.55	.23	.17	.064***

The table presents descriptive statistics from the sample. It presents for different cuts of the sample, based on the home country of the MNC, the total number of MNC firms, the number of subsidiaries, the proportion of those subsidiaries that are foreign (horizontal) subsidiaries, the average knowledge intensity of the foreign subsidiaries, the average knowledge intensity for the domestic subsidiaries, and the difference between these averages, denoted by Δ. Stars represent statistical significance of the difference: $^{*}p < 0.10, ^{**} p < 0.05, ^{***} p < 0.01$

Table 2: Determinants of Foreign Replication of Production

Dependent Variable: Horizontal Foreign Subsidiary Binary Variable			
	(1)	(2)	(3)
log(k)	-0.1610	-0.1600	-0.3427
	(0.096)*	(0.079)**	(0.090)***
log(t)	-0.0158	-0.0156	-0.0089
	(0.028)	(0.025)	(0.020)
GDpc (affiliate)	0.3271		
	(0.124)***		
Pop (affiliate)	-0.0663		
	(0.018)***		
Cap per worker (subsidiary)	-0.2772		
	(0.058)***		
HumCap per worker (subsidiary)	-0.5447		
	(0.198)***		
Downstream Linkage	0.1905	0.1951	0.2366
	(0.074)**	(0.078)**	(0.084)***
Upstream Linkage	0.5607	0.5159	0.4744
	(0.055)***	(0.055)***	(0.067)***
N	80542	80632	80487
Adj. R2	0.59	0.62	0.67
φ_h	Y	Y	Y
η_s	N	Y	-
$\eta_s \times \theta_P$	N	-	Y

The left hand side variable is a binary variable that takes the value 1 if the subsidiary is foreign. The variables in the right hand side include the unit shipping cost associated with the industry, knowledge intensity measures and other controls. All specifications include MNC fixed effects. Robust standard errors clustered at the industry and MNC level are presented in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table 3: Summary Statistics for Foreign Subsidiaries

Variable	**N**	**Mean**	**sd**	**Min**	**Max**
Distance, km (log)	25,166	8.201	1.15	1.1	9.9
Knowledge Intensity (log)	25,166	4.112	0.13	3.6	4.4
Unit Shipping Cost (log)	25,166	-1.894	0.76	-3.6	-0.1
Ratio GDP per capita (log)	25,151	0.125	0.71	-2.7	3.8
Ratio Population (log)	25,151	-0.423	1.71	-6.5	5.6
Ratio Capital per Worker (log)	25,151	0.350	0.98	-2.9	4.9
Ratio Human Capital (log)	25,151	0.118	0.33	-1.1	2.4
Upstream Linkage	25,166	0.290	0.45	0.0	1.0
Downstream Linkage	25,166	0.153	0.36	0.0	1.0
Ease of Communication Variables					
Non Stop Air Route	25,166	0.330	0.47	0.0	1.0
Working Hours Overlap	25,166	6.960	2.47	0.0	10.0

The table presents descriptive statistics from the sample of foreign subsidiaries, for distance in between the headquarters and the subsidiary, the knowledge intensity of the foreign subsidiary and the unit trade cost. The lower panel presents headquarters-subsidiary variables that measure the ease of communication in between them.

Table 4: Distance, Knowledge Intensity and Ease of Communication

Dependent Variable: Log Distance to Headquarters				
	(1)	(2)	(3)	(4)
log(k)	-0.5292	-0.3960	-1.4060	-1.2865
	(0.268)**	(0.230)*	(0.515)***	(0.477)***
log(k)XdirectFlightOAG	-0.0955	0.0414		
	(0.200)	(0.200)		
log(k)Xwhours_overlap			0.1520	0.1600
			(0.071)**	(0.071)**
directFlightOAG	-0.1020	-0.5741		
	(0.829)	(0.835)		
whours_overlap			-0.9359	-0.9742
			(0.304)***	(0.305)***
log(t)	-0.1091	-0.0381	-0.0632	-0.0276
	(0.079)	(0.054)	(0.052)	(0.035)
GDpc (affiliate)	-0.1282		-0.1277	
	(0.267)		(0.183)	
Pop (affiliate)	0.3237		0.1628	
	(0.043)***		(0.029)***	
Cap per worker (subsidiary)	-0.4883		-0.6073	
	(0.110)***		(0.095)***	
HumCap per worker (subsidiary)	1.7045		1.4679	
	(0.503)***		(0.281)***	
Downstream Linkage	0.0489	0.0721	-0.0714	-0.0081
	(0.099)	(0.069)	(0.076)	(0.037)
Upstream Linkage	0.0898	0.0740	0.0013	-0.0349
	(0.117)	(0.076)	(0.076)	(0.033)
N	25166	25237	25166	25237
Adj. R2	0.60	0.67	0.81	0.87
φ_h	Y	Y	Y	Y
η_s	N	Y	N	Y

The table presents results for the estimation of Specification (2) using a sample of foreign subsidiaries of MNCs, including controls and interactions with between log(k) and ease of communication proxies. The left hand side variable is the distance from the foreign subsidiary to its global headquarters in logs. The variables on the right hand side include knowledge intensity of the affiliate's economic activity (in logs), the unit shipping cost (in logs) and other controls. The right hand side also includes variables measuring the ease of communication between a headquarters and its subsidiaries, both on their own and interacted with log(k). All specifications include MNC fixed effects. Robust standard errors clustered at the industry and MNC level are presented in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table 5: Regression discontinuity estimation

Dependent Variable: Foreign subsidiary's knowledge intensity (log)				
	150	250	350	Optimal
closerTZ	0.0084	0.0060	0.0075	0.0066
	(0.003)**	(0.003)*	(0.002)***	(0.002)***
distTZzero	-0.0000	-0.0000	-0.0000	-0.0000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
closerTZ × distTZzero	-0.0000	0.0000	-0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
N	5232	9702	14179	12367
Adj. R2	0.87	0.87	0.86	0.87

The table presents results for the estimation of Specification (3) using a sample of foreign subsidiaries of MNCs estimated using several bandwidths for the running variable specified in each column. The last column uses the optimal bandwidth computed using the methodology described in Cattaneo et al. (2018) who build on the work by Imbens and Kalyanaraman (2012). The estimation uses a triangular weight scheme, giving higher weight to observations closer to the cutoff point. All specifications include MNC fixed effects and time zone line fixed effects. Robust standard errors clustered at the MNC and time zone line level are presented in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table 6: Regression discontinuity estimation for control variables

Dependent Variable: Control variables						
	$log(t_s)$	US_s	DS_s	$log(GDPpc_s)$	$log(Kpw_s)$	$log(Hpw_s)$
closerTZ	-0.0240	0.0246	0.0186	0.0028	-0.0359	-0.0114
	(0.018)	(0.017)	(0.012)	(0.028)	(0.052)	(0.022)
distTZzero	0.0001	-0.0001	-0.0001	-0.0003	-0.0001	-0.0001
	(0.000)*	(0.000)**	(0.000)*	(0.000)**	(0.000)	(0.000)**
closerTZ × distTZzero	-0.0001	0.0002	-0.0000	0.0002	-0.0003	-0.0000
	(0.000)*	(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)
N	12006	12367	12367	12309	12323	12323
Adj. R2	0.92	0.58	0.74	0.91	0.93	0.92

The table presents results for the estimation of Specification (2) using a sample of foreign subsidiaries of MNCs, including controls and interactions with between log(k) and ease of communication proxies. The left hand side variable is the distance from the foreign subsidiary to its global headquarters in logs. The variables on the right hand side include knowledge intensity of the affiliate's economic activity (in logs), the unit shipping cost (in logs) and other controls. The right hand side also includes variables measuring the ease of communication between a headquarters and its subsidiaries, both on their own and interacted with log(k). All specifications include MNC fixed effects. Robust standard errors clustered at the industry and MNC level are presented in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Appendix for
The hardships of long distance relationships: time zone proximity and knowledge transmission within multinational firms

May 30, 2018

A Conceptual Framework

In this section I augment the model by Helpman et al. (2004) – referred to as HMY hereafter – by including a new parameter capturing the intra-firm cost of transmitting knowledge between headquarters and foreign subsidiaries. This extension allows us to understand how the cost of knowledge transmission faced by firms affects their decision to serve foreign markets. First the common set-up is described and then the proper adaptation is incorporated.

As in HMY, there are N countries producing H+1 sectors with labor as the only input of production. H sectors (indexed 1, 2,...,H) produce a differentiated good, while the other sector (indexed 0) produces an homogenous good (which serves as the numeraire). In any given country, individuals spend a share $\beta_h > 0$ of their income on sector h, such that $\sum\limits_{0 \leq h \leq H} \beta_h = 1$. Country i is endowed with L^i units of labor and the wage rate in this country is denoted by w^i.

Consider now a particular differentiated sector, h. For simplicity of notation, the index h is dropped in the next few paragraphs, but it is implicit that all sector specific variables may vary across sectors.[29] In order to enter the industry in country i a firm bears a fixed and sunk cost f_E denominated in units of labor. After bearing this cost, the potential entrant learns its labor-per-unit cost, a, drawn from a common and known distribution $G(a)$. Upon observing this cost, the firm may choose not to enter, and thus bear no additional costs and receive no revenues. If it chooses to produce, however, an additional cost of f_D units of labor is incurred. There are no other fixed costs if the firm chooses to produce and sell in the local market only.

The firm can choose to serve a foreign market either by exporting or creating a foreign subsidiary. If the firm chooses to export, it bears an additional cost of f_X (per country it exports to). If it chooses to create a foreign affiliate, it incurs an additional cost of f_I for every foreign market it chooses to serve this way. Similar to HMY, f_X can be interpreted as the cost of forming a local distribution and service network in the foreign market, and f_I includes all of these costs, as well as the cost of forming a subsidiary in the foreign country and the overhead production costs embodied in f_D. Hence, $f_I > f_X > f_D$.

[29] Some sector-specific variables are explicitly kept in the notation, such as t and k, since these variables will be relevant in the empirical analysis.

The homogenous good is freely traded at no cost.[30] Differentiated goods that are exported from country i to country j are subject to a "melting-iceberg" transport cost $\tau(t, d_{ij})$ which is an increasing function of the per unit shipping cost of the good (denoted by t, and proxies for weight or other good specific characteristics) and the distance between countries i and j (denoted by d_{ij}). It is assumed that that $\tau(t, d_{ij}) > 1$. That is, a firm in country i has to ship τ units of a good for 1 unit to arrive in country j.

Analogously, serving a foreign market through an affiliate is subject to a marginal cost $\kappa(k, d_{ij})$ related to the transfer of knowledge. $\kappa(k, d_{ij})$ is assumed to be an increasing function of both the knowledge intensity of the good (represented by k) and distance (d_{ij}). The cost of transferring knowledge includes resources and time used for communicating with foreign affiliates to transmit proper knowledge required for efficient production. It is assumed that $\tau(t, d_{ij}) > \kappa(k, d_{ij}) > 1$ for all goods. The last inequality implies that for a multinational corporation, the cost of selling 1 unit of a good through a foreign affiliate is $\kappa(k, d_{ij})$.

The cost of knowledge transmission in knowledge intensive sectors being higher is justified given that these sectors require higher interaction and communication among their workers. Thus, firms pay for business travel and communication services that occur more often within these sectors. In addition, and perhaps more importantly, knowledge intensive activities usually encompass tasks with higher probability of failure and thus requiring trained and experienced workers. This too raises operational costs. Assuming that knowledge transmission costs are increasing in distance is consistent with empirical evidence (e.g., Jaffe et al., 1993; Bottazzi and Peri, 2003; Keller, 2002, 2004; Bahar et al., 2014), and a standard way to model these costs in the literature studying MNCs (e.g., Ramondo and Rodriguez-Clare, 2013; Ramondo, 2014; Arkolakis et al., 2013; Tintelnot, 2014; Keller and Yeaple, 2013; Gumpert, 2015).

All the producers which serve a market engage in monopolistic competition. Consumer preferences across varieties of a differentiated product h have the standard CES form, with an elasticity of substitution $\varepsilon = \frac{1}{1-\alpha} > 1$. It is well known that these preferences generate a demand function $A^i p^{-\varepsilon}$ for each product in the industry in country i, where $A^i = \frac{\beta}{\int_0^{n^i} p^i(s)^{1-\varepsilon} ds} E^i$, n^i is the measure of firms active in the industry in country i, and $p^i(s)$ is the consumer price for a product indexed s.

In this setting, an active producer with labor requirement of a optimally sets a price of $\frac{w^i a}{\alpha}$. Consequently, the price of a locally produced good is $\frac{w^i a}{\alpha}$, the price of a good which is exported to country j is $\frac{\tau(t,d_{ij}) w^j a}{\alpha}$, and the price of a good that is sold by a foreign affiliate in country j is $\frac{\kappa(k,d_{ij}) w^j a}{\alpha}$, where a is the labor required for the producer to manufacture one unit of the product.

In what follows, it is shown that the balance of forces ruling the trade-off of serving a foreign market through exports or FDI is influenced by the knowledge

[30] Thus, as long as the numeraire good is produced in all countries the wage rate is equalized.

intensity of the product.

The assumption that the numeraire good is produced in each country simplifies the analysis, as it implies that the wage rate is equalized across all countries and is equal to 1. Hence, the operating profit for a firm in country i with a labor coefficient of a from serving the domestic market maybe expressed as $\pi^i_D = a^{1-\varepsilon}B^i - f_D$, where $B^i = (1-\alpha)\frac{A^i}{\alpha^{1-\varepsilon}}$. The additional profits from exporting to country j are $\pi^i_X = (\tau(t,d_{ij})\cdot a)^{1-\varepsilon}B^j - f_X$ and those from selling in country j through a foreign affiliate are $\pi^i_I = (\kappa(k,d_{ij})\cdot a)^{1-\varepsilon}B^j - f_I$. B^i represents demand parameters for country i and are considered exogenous to each individual firm.

Hence, in this setting, the productivity parameter a will be critical for a firm's decision of whether to serve the local market only or to serve foreign markets, either through exports or FDI. The sorting pattern is similar to the one in HMY and is based in the following equations:

$$(a_D)^{1-\varepsilon}\cdot B^i = f_D, \qquad \forall i \tag{A1}$$

$$(\tau(t,d_{ij})\cdot a_X)^{1-\varepsilon}\cdot B^j = f_X, \quad \forall i, \forall j \neq i \tag{A2}$$

$$\left[\kappa(k,d_{ij})^{1-\varepsilon} - \tau(t,d_{ij})^{1-\varepsilon}\right]\cdot a_I^{1-\varepsilon}\cdot B^j = f_I - f_X, \quad \forall i, \forall j \neq i \tag{A3}$$

Similar to HMY, the first two equations define the productivity thresholds after which firms will sell domestically or export, respectively. The minimum productivity threshold after which firms will engage in FDI is derived from Equation (A3).[31] This threshold is defined as:

$$a_I^{1-\varepsilon} = \frac{f_I - f_X}{\left[(\kappa(k,d_{ij}))^{1-\varepsilon} - (\tau(t,d_{ij}))^{1-\varepsilon}\right]B^j}, \forall i, \forall j \neq i \tag{A4}$$

Predictions derived from this model will serve as the basis for the empirical analysis. The implications of the original HMY model are straightforward. An increase in $\tau(t,d_{ij})$, either through an increase in either t or d_{ij}, will result in lower π_E making it more likely to substitute exports with FDI. This is part of the mechanism of the concentration-proximity trade-off. However, with the inclusion of $\kappa(k,d)$ into the model, some new predictions arise, assuming full symmetry in fixed costs and demand variables for all sectors and countries. The propositions are presented in terms of $\phi(a_I) = a_I^{1-\varepsilon}$.

Proposition 1. *For a higher level of k fewer firms will substitute exports towards FDI.*

Proposition (1) is a direct consequence of adding κ into the model. All else equal, for higher levels of k, the threshold $a_I^{1-\varepsilon}$ for which above it FDI is more profitable than exports is larger, implying FDI will be less likely for

[31] Condition (A3) will have a positive solution if we assume $\kappa(k,d_{ij})^{\varepsilon-1}f_I > \tau(t,d_{ij})^{\varepsilon-1}f_X > f_D$, which is homologous to condition (1) in HMY (with equal wages across countries), but including κ.

sectors with higher k. The graphical representation of the model in Figure A1 shows the case for two sectors that differ in their knowledge intensity, $\underline{k}$ and $\overline{k}$ (where $\overline{k} > \underline{k}$). Notice that the profit functions for both sectors originate in the same fixed cost value f_I, but the function is flatter for the sector $\overline{k}$ (dashed line). Hence, the productivity threshold required for a firm to substitute exports with FDI becomes higher for sectors with higher levels of k. That is, $(a_I^{ij,\overline{k}})^{1-\varepsilon} > (a_I^{ij,\underline{k}})^{1-\varepsilon}$.

[Figure A1 about here.]

Proposition 2. *Conditional on expansion, firms face a trade-off between k and d*

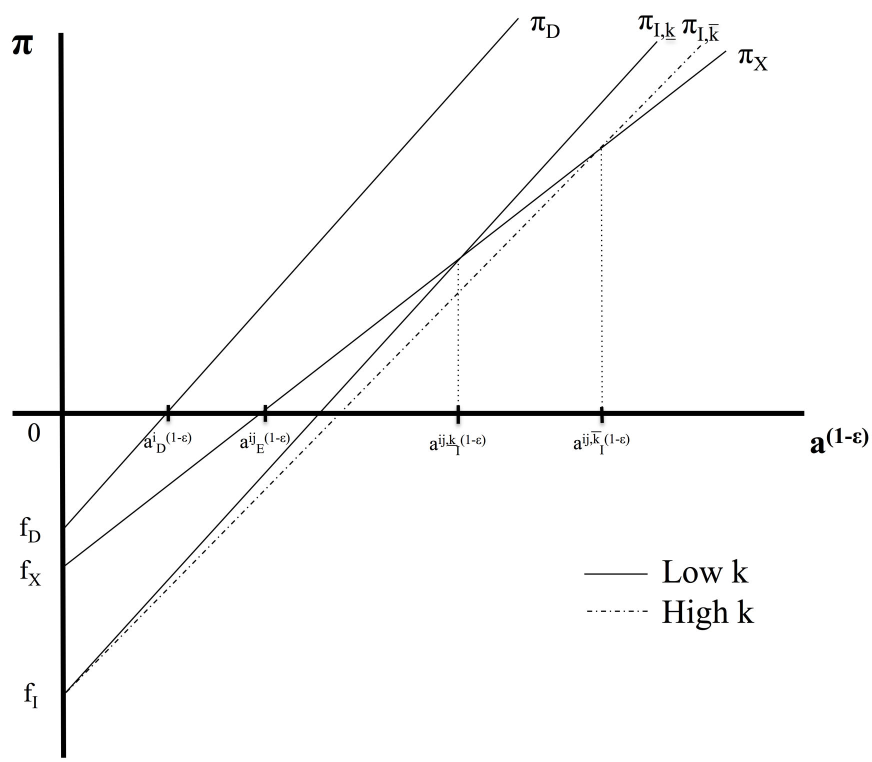
Proposition (2) follows from the assumptions that $\partial\kappa/\partial k > 0$ $\partial\kappa/\partial d > 0$. Figure A2 exemplifies this proposition.[32] The figure represents the isoprofit curves of a firm for both exports and FDI activities as a function of k and d. A firm will choose to expand horizontally (over exporting) in every combination of k and d where $\pi_D + \pi_I > \pi_D + \pi_E$, consistently with the equations described above. Notice that conditional on FDI, for a given level of profits $\pi_D + \pi_I$, the model predicts a trade-off between k and d.

[Figure A2 about here.]

The predictions coming out of the model following the inclusion of an intra-firm cost of transferring knowledge (κ) have testable implications in the data. First, *ceteris paribus*, industries with higher levels of knowledge intensity will be *less likely* to expand horizontally to foreign destinations. Second, conditional on horizontal expansion, firms will face a trade-off between distance to the headquarters and knowledge intensity of the activity.

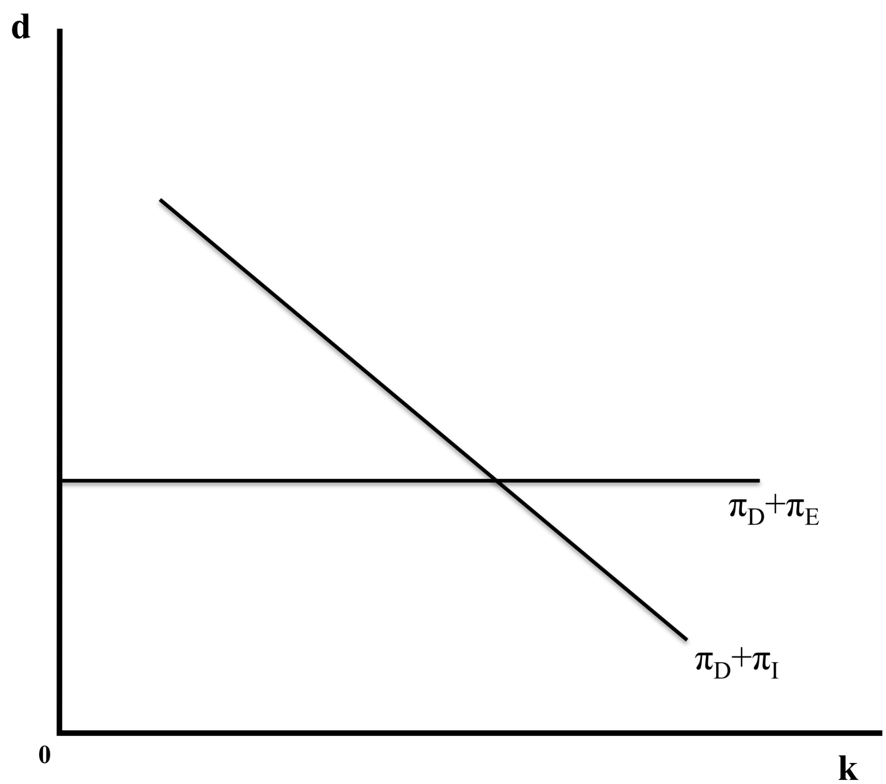
[32]The exact shape of the curves will depend on assumptions of $\kappa(k, d)$. I assume a general form for now.

Figure A1: Increase in k (knowledge intensity)



Graphical representation of the model, for a case considering two sectors with different levels of k, where $\overline{k} > \underline{k}$. The result suggests that the threshold a_I is an increasing function of k. Thus, FDI will be less likely for sectors with higher k.

Figure A2: Profit curves in k and d dimension



The figure is a graphical representation of a firm's profit as a function of k and d. The curve $\pi_D + \pi_E$ represents total profits for an exporting firm, while the curve $\pi_D + \pi_I$ represent total profits for a firm engaging in FDI instead.

B On the Dun & Bradstreet Dataset

B.1 Reliability

The Worldbase dataset collected by Dun & Bradstreet is sourced from a number of reliable organizations all over the world, including public registries. According to Dun & Bradstreet's website, "the data undergoes a thorough quality assurance process to ensure that our customers receive the most up-to-date and comprehensive data available".[33] However, it is important to acknowledge that, given the lack of access to public registries for every country, it is not possible to asses with full accuracy the representativeness of the data. Alfaro and Charlton (2009), however, compare the dataset with the US multinational firms sample by the US Bureau of Economic Analysis, and find consistencies between the two datasets.

Some basic relationships drawn from the sample behave as expected. For instance, the number of countries in which an MNC has foreign affiliates is related to the overall size of the MNC, as can be seen in Figure B1. In particular, the figure shows the relationship between the size of MNC firms (in number of establishments in the left panel, and in total number of employees in the right panel[34]) against the number of foreign countries in which their subsidiaries are located (on the vertical axis). Each observation in the scatterplot is an MNC labeled with its headquarters' country ISO3 code. The figure shows smaller MNCs are present in fewer countries, while larger MNCs tend to be more spread out in terms of the number of countries they have a presence in.

[Figure B1 about here.]

Focusing the analysis on the within-MNC dimension, by adding MNC fixed effects to all specifications, significantly diminishes the sampling concerns further. This is because, while methods for gathering information may not be symmetric across countries, they would not systematically differ by firm. It is reasonable to assume that the per-country likelihood of missing data would be the same for all firms, controlling for the location of the headquarters of the MNC. Thus, concerns regarding biases caused by possible sampling asymmetries are not particularly large for the purpose of this empirical exercise.

B.2 Industries

While the dataset has information on up to six industries per plant (a main one plus five other) the number of establishments that report more than one activity varies dramatically per country. The left panel of Figure B2 shows the average number of reported industries across all subsidiaries per country, while the right panel shows, per country, the percentage of firms reporting one, two, three, four, five or six industries. In most countries, the average number of reported firms

[33] http://dnb.com.au/Credit_Reporting/The_quality_of_DandBs_data/index.aspx
[34] Including their domestic plants for both.

is below two; and the majority of firms in more than half the countries report only one SIC code.

[Figure B2 about here.]

I also present results of the distribution of sectors among foreign affiliates, to understand whether in the sample there are some sectors that are more likely to appear (i.e. be reported) than others. In terms of industries, the distribution of different sectors in the sample is not homogenous, as can be seen in Figure B3. Some sectors are more prevalent than others in the data. The industries that appear the most in the data are Ready-Mixed Concrete (SIC 3273), Pharmaceutical Preparations (SIC 2834) and Motor Vehicles Parts (SIC 3714). To alleviate concerns on how this distribution could affect the results, all the standard deviations calculations allow for clustering at the industry level.

[Figure B3 about here.]

In addition, it is worth emphasizing that each foreign subsidiary in the sample manufactures a specific product. Hence, if a MNC has several foreign subsidiaries, then each one of those could be manufacturing a different product (in its 4 digit classification). The sample that a single MNC that has more than one foreign subsidiary could be manufacturing more than one product. Figure B4 shows that larger MNCs (as measured by number of affiliates) tend to make a larger number of different products.

[Figure B4 about here.]

B.3 Using the input-output table to define vertical relationships

In order to filter out from the definition of horizontal those links that could also be defined as vertical, either upstream or downstream, I use the US input-output provided by Fan and Lang (2000). I follow the methodology suggested by Alfaro and Charlton (2009) and Acemoglu et al. (2009) to define vertical relationships.

More in general, the diagram in Figure B5 is useful to understand how horizontal and vertical links are defined in the dataset. Within a single MNC firm, an horizontal link is defined as a foreign subsidiary that is classified under the same SIC code as any of its domestic subsidiaries. Then I use the US I/O table by Fan & Lang (2000) to define vertical relationships, both downstream and upstream. A subsidiary is defined as upstream vertical if its main economic activity is an input of $0.05 or more per each dollar of output of any of the domestic subsidiaries of the firm. Similarly, a subsidiary is defined as downstream vertical if any of the domestic subsidiary provides an input to it of $0.05 or more per each dollar of output.

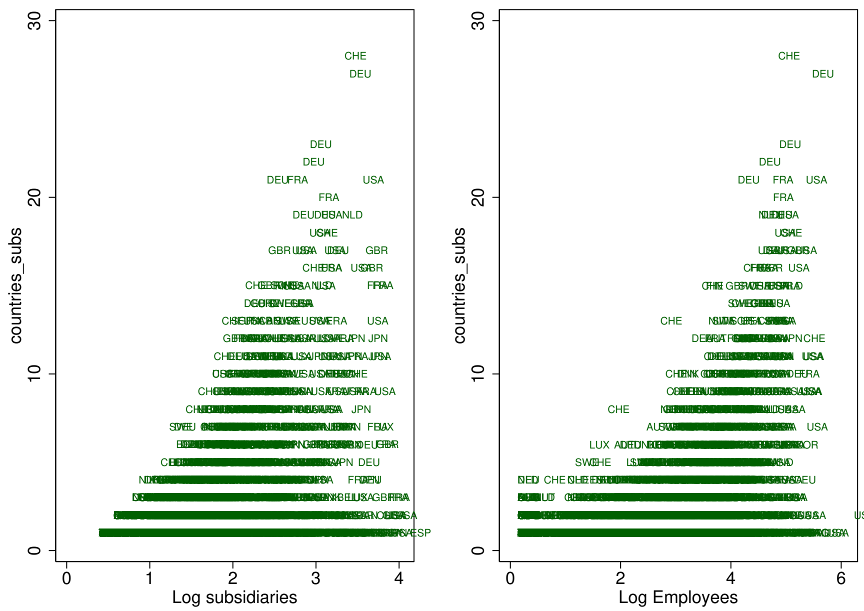
After such classification, those subsidiaries that fall into both categories (horizontal and vertical) are filtered out from the horizontal classification. This implies that the sample classifies as horizontal only final goods (which is the matter of study of the theoretical framework presented in Online Appendix Section A).

[Figure B5 about here.]

The use of $0.05 in the main body of the paper follows the precedent set by Alfaro & Charlton (2009), but its choice is not determinant for the results of the paper.

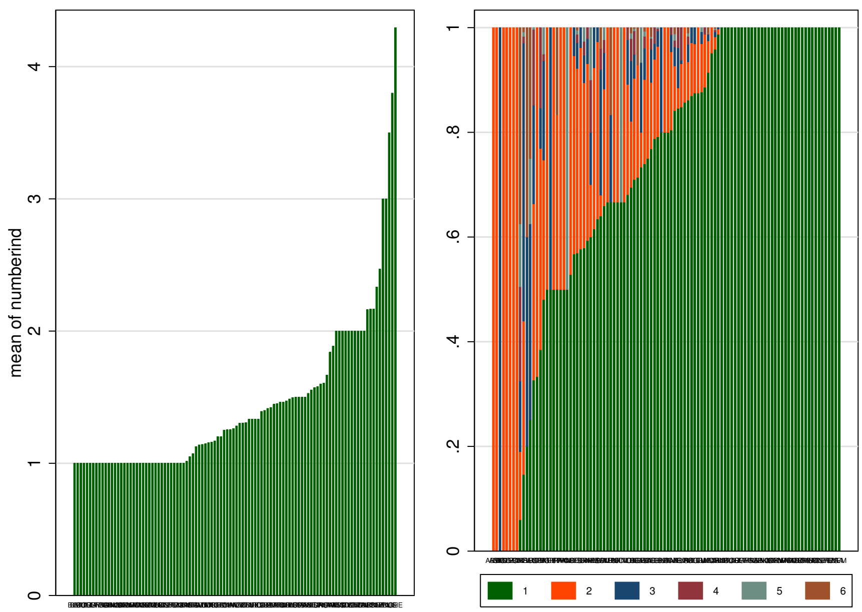
A limitation of this methodology is that technologies might vary across countries, and hence, the US I/O table would loss some validity in defining upstream or downstream relationships. While acknowledging this limitation I assume that the US I/O table is a good proxy for measuring vertical links, regardless of the country, in line with the previous literature.

Figure B1: MNC size vs. number of countries



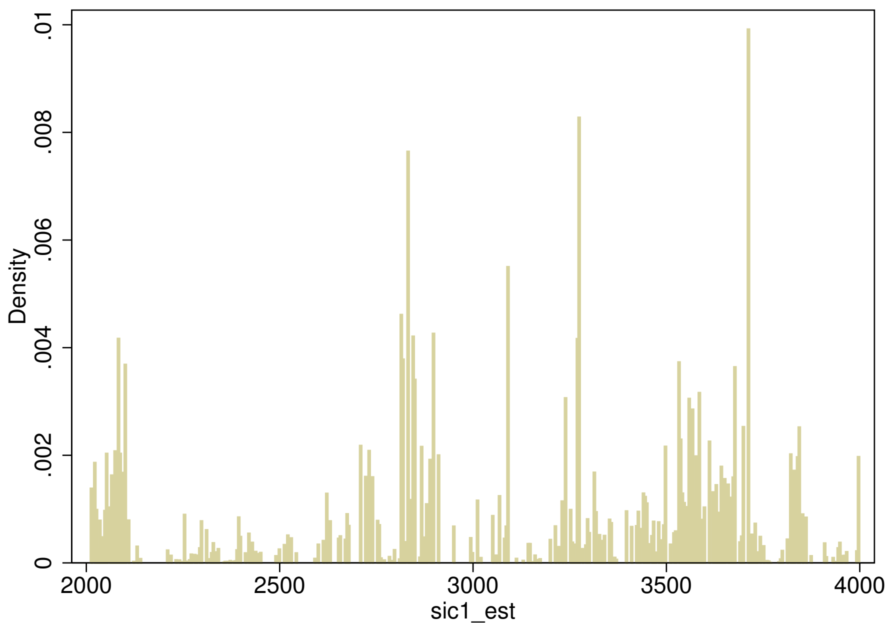
The figure shows the relationship between the size of MNC (horizontal axis) and the number of foreign countries they are active in (vertical axis). In the scatterplots, each observation is an MNC, labeled with the ISO3 code of the country where its headquarters is located. The left panel measures the firms' size by the total number of subsidiaries it has (both domestic and foreign), while the right panel uses the total employees (both in domestic and foreign plants).

Figure B2: Distribution of reported SIC codes by plant, per country



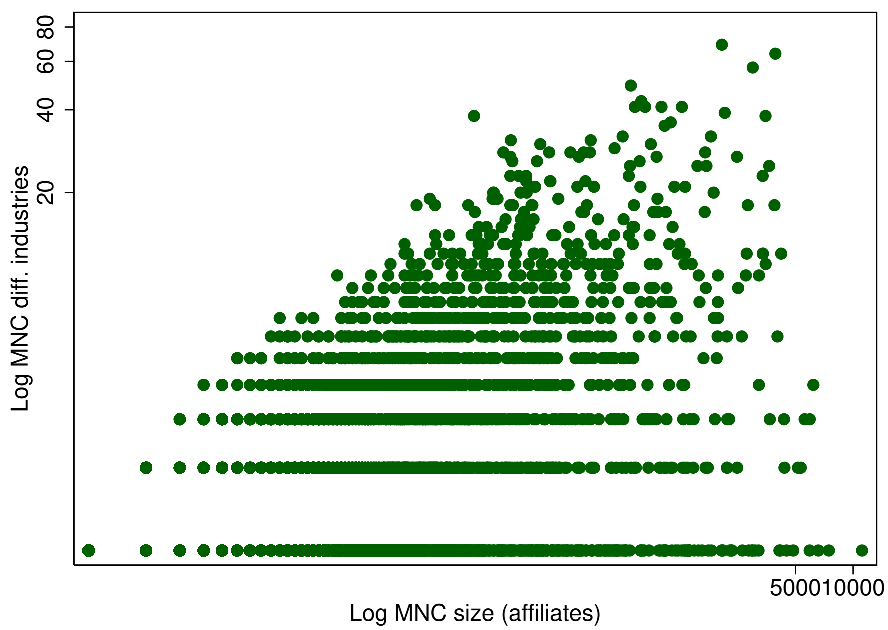
The figure describe the distribution of number of industries reported by establishment in the sample. The left panel shows the average number of reported industries across all subsidiaries per country, while the right panel shows, per country, the percentage of firms reporting one, two, three, four, five or six industries.

Figure B3: Histogram of SIC codes in the sample



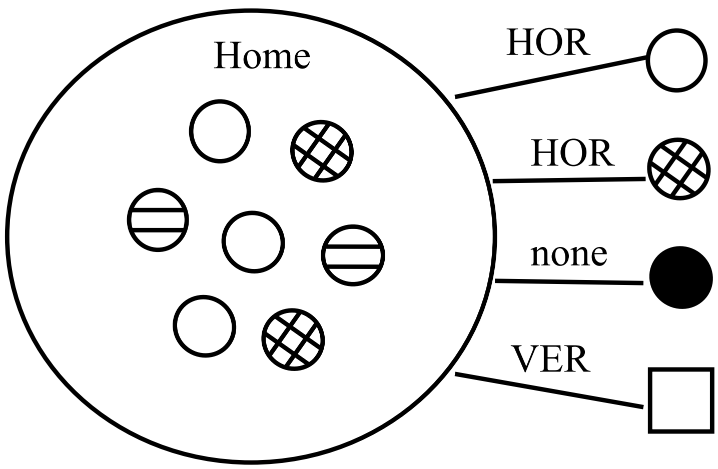
The figure is an histogram of the SIC industries reported in the dataset. Each bin represents the frequency of a particular SIC code within the manufacturing sector. Notice that the SIC classification is not fully continuos, what explains the zero values in the figure.

Figure B4: Number of different industries Vs. MNC size



The figure plots the relationship between MNC size and total number of (different) industries the MNC is active in through its foreign affiliates. The figure reveals that larger MNCs (measured in terms of number of subsidiaries) tend to make a larger number of different products.

Figure B5: Definition of Horizontal and Vertical



The diagram describes the methodology used to classify foreign subsidiaries as horizontal expansions based on their reported economic activity vis-a-vis the economic activity of the MNC in its home country.

C O*NET knowledge intensity measures

This section describes characteristics of the knowledge intensity measures based on Bahar (2018). Figure C1 presents the distribution of the knowledge intensity measure used in the paper: experience plus training (based on experience plus on-site and on-the-job training figures of workers in each industry). As opposed to the R&D investment based variables used in the literature, the distribution of the O*NET based variables is smoother, and behaves more like a normal probability density function. Figure C2 presents the same graphs limiting the sample to manufacturing industries only.

[Figure C1 about here.]

[Figure C2 about here.]

Tables C1 presents the top and bottom ten products in the manufacturing division (SIC codes 2000 to 3999) ranked by the knowledge intensity measure.

[Table C1 about here.]

C.1 Advantage over R&D Measures

I find this measure correlates positively with other (proxies of) knowledge intensity measures used in the literature, such as the average R&D share of sales per industry. Correlation coefficient is 0.13 with R&D intensity computed using the Compustat dataset and compiled by (Keller and Yeaple, 2013) and 0.20 with R&D intensity computed using the Orbis dataset and compiled by Nunn and Trefler (2008), using manufacturing NAICS 4-digit industries.[35]

The R&D based measures, however, have three main shortcomings that could generate significant biases. First, these measures are concentrated in lower values and have a long tail, as shown in Figure C3. This is because most firms within those industries have no R&D investment whatsoever. For these industries in the lower end of the distribution, the intensity of their knowledge is almost indistinguishable. Second, since these measures are computed by averaging across each industry the R&D share of sales reported by a (random or not) sample of firms, they are likely to favor industries in which larger firms are more prevalent. This might happen in industries for which the barriers to entry are higher, and not necessarily knowledge intensive industries. Third, R&D investment might not be equally accounted for across all industries.

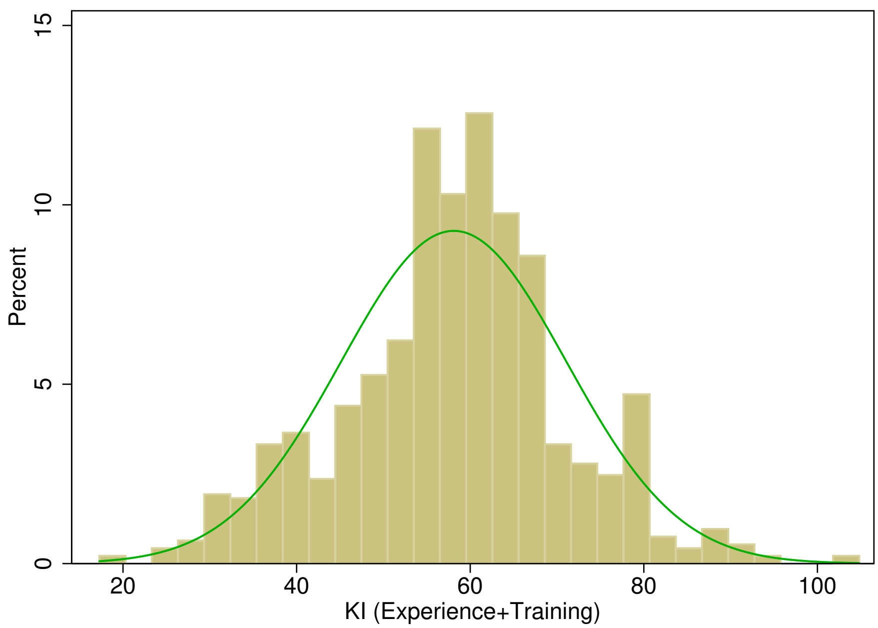
[Figure C3 about here.]

[35]It also correlate positively with other less popular measures that could proxy for knowledge intensity or complexity. The correlation coefficient with the share of non-production workers in total employment, from the NBER-CES Manufacturing Industry Database (Becker el al. 2013), is 0.68. Similarly, the correlation coefficient with the Product Complexity Index, developed by Hausmann el al. (2011), is 0.49. These correlations were computed using SIC 4 digits codes.

The O*NET based measures solve these issues. As shown in Figure C1, their distribution is smoother and closer to a two-tailed normal distribution.[36] Moreover, our measures do not rely on a sampling of particular firms, and they use the same standardized measure for all industries.

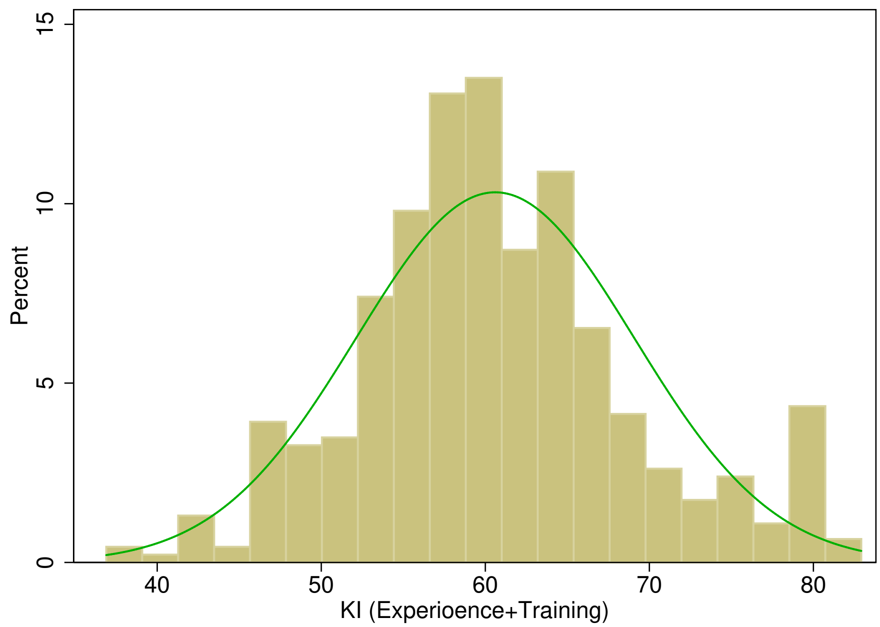
[36]According to the Skewness/Kurtosis test for normality, we cannot reject the hypothesis that these measures are normally distributed.

Figure C1: Histogram O*NET-based KI (All Industries)



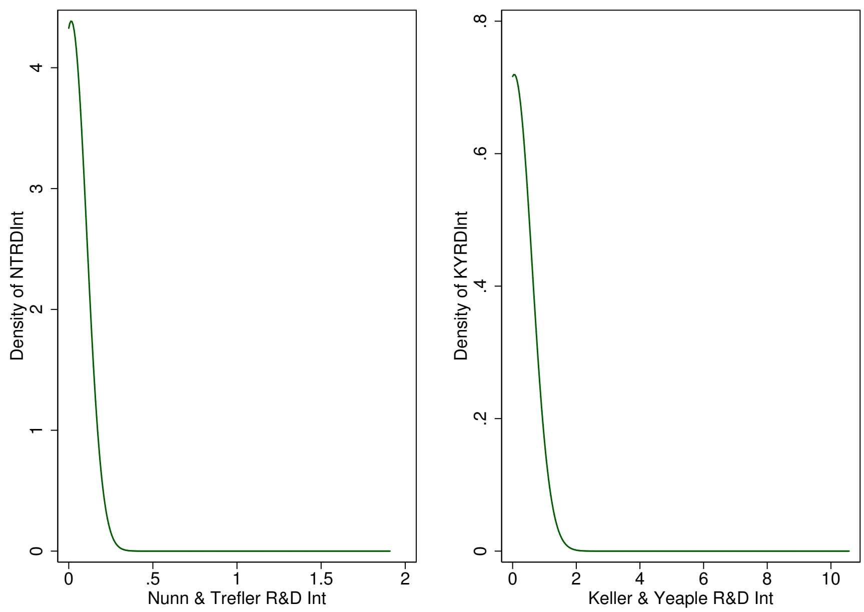
The figure shows the fitted distribution for the computed "experience plus training" O*NET-based knowledge intensity measures for all industries. Industries are defined in SIC 1987 4-digit industries.

Figure C2: Histogram O*NET-based KI (Manufacturing Only)



The figure shows the fitted distribution for the computed "experience plus training" O*NET-based knowledge intensity measures for manufacturing industries only. Industries are defined in SIC 1987 4-digit industries.

Figure C3: Distribution R&D Measures



The figure shows the fitted distribution for the R&D Intensity measures used in the literature. The left panel corresponds to the industry-specific R&D intensity computed using the Compustat dataset and compiled by Keller and Yeaple (2013) and the right panel corresponds to the same measure computed using the Orbis dataset and compiled by Nunn and Trefler (2008). Industries are defined in NAICS 4-digit industries.

Table C1: Top and bottom 10 manufacturing products, ranked by KI

Rank	SIC	Name	Value
		Ranking by Experience + Training, Top 10	
1	3669	Communications Equipment, NEC	82.92
2	3663	Radio and Television Broadcasting and Communications Equipment	82.92
3	3661	Telephone and Telegraph Apparatus (except consumer external modems)	81.45
4	3677	Electronic Coils, Transformers, and Other Inductors	79.97
5	3676	Electronic Resistors	79.97
6	3678	Electronic Connectors	79.97
7	3675	Electronic Capacitors	79.97
8	3671	Electron Tubes	79.97
9	3672	Printed Circuit Boards	79.97
10	3674	Semiconductors and Related Devices	79.97
		Ranking by Experience + Training, Bottom 10	
459	2013	Sausages and Other Prepared Meat Products (except lard made from purchased materials)	36.89
458	2011	Meat Packing Plants	36.89
457	2411	Logging	39.79
456	2077	Animal and Marine Fats and Oils (animal fats and oils)	41.39
455	2053	Frozen Bakery Products, Except Bread	41.53
454	2045	Prepared Flour Mixes and Doughs	41.53
453	2098	Macaroni, Spaghetti, Vermicelli and Noodles	41.53
452	2051	Bread and Other Bakery Products, Except Cookies and Crackers	41.53
451	2015	Poultry Slaughtering and Processing (poultry slaughtering and processing)	42.72
450	2052	Cookies and Crackers (unleavened bread and soft pretzels)	45.04

The table presents the top and bottom 10 manufacturing sectors ranked by the "experience plus training" O*NET based knowledge intensity measure.

D Proxying for demand: knowledge transmission in the proximity-concentration trade-off

The results on the determinants of horizontal expansion are not fully satisfactory since they do not account for the demand factor (i.e. it is not possible to see the all the locations where the MNCs faced a trade-off between exports and foreign affiliates and decided for the former). However, in order to explore for the role of knowledge intensity in the likelihood of horizontal expansion we follow the guidelines of Helpman, Melitz and Yeaple (2004) using BEA data on American MNCs for years 1999-2011 to show that sales of foreign affiliates decreases with the knowledge intensity level of the industry (using my own measures of knowledge intensity), after controlling for export volume for those industries. That is, by including exports in the specification I aim to control for sector specific demand of American exporters in that location. Thus, the specification I estimate is:

$$log(ForeignSales_{s,y}) = \beta_k log(k_s) + \beta_t log(t_s) + log(Exports_{s,y}) + \alpha_y + \varepsilon_{s,y} \quad \text{(D1)}$$

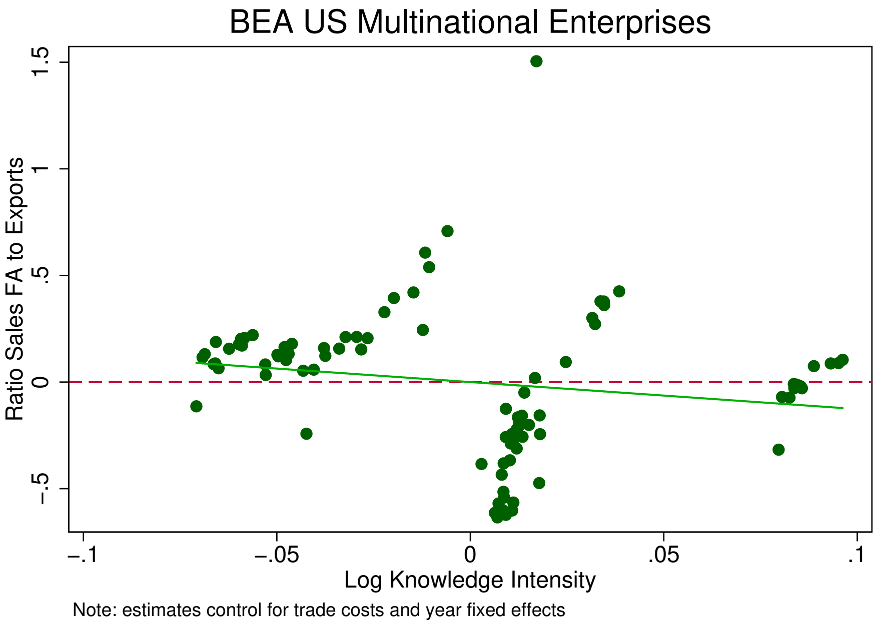
Where s indexes for industry and y for year. The dependent variable is the sales of foreign affiliates of US multinationals in the rest of the world for sector s, in millions of dollars. The right hand side includes the knowledge intensity (denoted by k) and the unit shipping cost (denoted by t) of sector s. It also includes US export volumes of industry s to the world in millions of dollars. There are in total seven sectors, and each is defined as a 3-digit NAICS code. The specification also includes year dummies (denoted by α_y). The results of this estimation are presented in Table D1.

[Table D1 about here.]

Column 1 estimates a linear regression where the dependent variable is the sales of foreign affiliates (in logs), controlling for total exports to the same destination in the same industry code. Column 2 estimates a linear regression where the dependent variable is the ratio of sales of foreign affiliates to exports, similarly to Helpman, Melitz and Yeaple (2004). Column 3 replicates column 2 but excludes outliers in the sales to exports distribution. It can be seen that the estimator for β_k is negative and statistically significant, regardless of whether the specification uses the total sales of foreign affiliates controlling for exports (column 1), or the ratio of foreign sales to exports (column 2 and 3). Similarly, β_t is estimated to be positive and significant, as expected (i.e. industries with larger trade cost will generate incentives for firms to create foreign affiliates abroad to substitute for exports). Figure D1 plots the estimation of β_k in the first column of Table D1.

[Figure D1 about here.]

Figure D1: Sales of Foreign Affiliates for American MNCs and Knowledge Intensity



The vertical axis is the log-sales of foreign affiliates and the horizontal axis is the log knowledge intensity measure.

Table D1: Sales of Foreign Affiliates for American MNCs and Knowledge Intensity

Dependent Variable: Log Sales of Foreign Affiliates			
	log(SalesFA)	Ratio	Ratio (no outliers)
log(k)	-1.2664	-1.5221	-1.0200
	(0.292)***	(0.650)**	(0.298)***
log(t)	0.7290	1.6101	1.7935
	(0.104)***	(0.229)***	(0.095)***
log(exp)	1.1858		
	(0.048)***		

The table presents results for the estimation of Specification (D1) using the Bureau of Economic Activity's data on American MNCs for years 1999-2011. The left hand side variable is the log of foreign sales of American MNCs in each 3-digit NAICS code (column 1) or the ratio of foreign sales to exports (columns 2 and 3). Column 3 exclude outlier observations in terms of the ratio. The variables in the right hand side include the unit shipping cost associated with the industry, a knowledge intensity measure and the exports of the US to the rest of the world in each 3-digits NAICS category (column 1 only). All specifications include year fixed effects. Robust standard errors clustered at the year level are presented in parentheses.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

E Regression discontinuity uniform weighting scheme

Table E1 repeats the estimation of Specification (3), using a uniform weighting scheme which provides equal weight to all observations within the selected bandwidth. Results are robust to those presented in Table 5 in the main body of the paper.

[Table E1 about here.]

Table E1: Regression discontinuity estimation, uniform weighting scheme

Dependent Variable: Foreign subsidiary's knowledge intensity (log)				
	150	250	350	Optimal
closerTZ	0.0084	0.0060	0.0075	0.0066
	(0.003)**	(0.003)*	(0.002)***	(0.002)***
distTZzero	-0.0000	-0.0000	-0.0000	-0.0000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
closerTZ × distTZzero	-0.0000	0.0000	-0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)
N	5232	9702	14179	12367
Adj. R2	0.87	0.87	0.86	0.87

The table presents results for the estimation of Specification (3) using a sample of foreign subsidiaries of MNCs estimated using several bandwidths for the running variable specified in each column. The last column uses the optimal bandwidth computed using the methodology described in Cattaneo et al. (2018) who build on the work by Imbens and Kalyanaraman (2012). The estimation uses a uniform weight scheme, giving same weight to all observations within the bandwidth. All specifications include MNC fixed effects and time zone line fixed effects. Robust standard errors clustered at the MNC and time zone line level are presented in parentheses.

$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$