Multinational Firm's Location Decision: Distance vs. Connectivity

Trang Tran*

Prepared for The 78th JSIE Annual Meeting on September 29, 2019 @ IDE-JETRO

Abstract

Using Japanese foreign and domestic affiliates' data, we examine how the inter-relationship within firms and groups of firms affects the firm's decision on the location of a new affiliate. We find that in general, the presence of the firm or group's prior affiliate increases the possibility of the firm having a new affiliate in the region. However, the type of affiliate matters in the sense that firms cluster their affiliates by sectors. In addition, changing the size of regions does not affect the results quantitatively but shows different patterns among affiliates' sectors.

JEL Classification: F23,L22,R3
Keywords: Firm Location, Multinational Firms, Vertical links

1 Introduction

The determinants of Foreign Direct Investment (FDI) locations have been of great interest not only to researchers but also to politicians from both developed and developing countries. With the increasing availability of detailed data, researchers have been able to look into the complexity of multinational firms' international location strategies. Most firms are now doing both horizontal and vertical FDI, export-platform FDI or hybrid FDI (Antras and Yeaple, 2015). The decision on where to place foreign affiliates as well as what types of affiliates to open has been studied extensively in both literature of FDI locations and New Economic Geography in terms of market access, supplier access, trade costs and factor costs of entering the market. (See Head and Mayer (2004), Duranton and Puga (2005),

*Graduate School of Economics, Hitotsubashi University

Blonigen et al. (2007), Defever (2012) for recent literature.) In particular, spatial agglomeration has been found not only across-firm within industries (Head et al. (1995), Strauss-Kahn and Vives (2009) for example) but also within-firm (Mayer et al. (2010) and Defever (2012) for example). There are also many studies on global city networks which emphasize the connectivity among cities through firms' flows of services (Taylor, 2001). With this connectivity among cities, firms can overcome spatial distances to place their headquarters in different regions (Belderbos et al., 2015).

In this paper, we examine these effects on the affiliates' location choice of manufacturing firms by building indices that reflect the agglomeration effects within-firm and within-financial groups of firms[1], and the importance of a region in Japanese financial, services, transportation and wholesale networks. Our contribution is to combine these two types of ideas in a framework that allows us to compare the sizes of these effects.

Our paper is different from Mayer et al. (2010) and Defever (2012) in the scope of study and the construction of relevant indices. The former examines French firms investing abroad where within-firm connection is defined as belonging to the same financial group. They found this contributes to a large portion of French firms locating affiliates domestically rather than investing abroad. Our paper focused solely in location decision of foreign affiliates. The latter investigates the case of foreign firms investing in the European Union (the EU) and focuses on the distance between affiliates of different functions. Our paper, while including interactions among affiliates of different functions, separates the industry-level, financial-group level and firm-level agglomeration effects.

Differing from Belderbos et al. (2017) regarding city connection index, we examine not only headquarters but also other functions of affiliates. In addition, they use a composite measure of connectivity based on producer services, airport passenger flows, and the intensity of international co-invention. We use activities of Japanese firms in different sectors. We should point out that our paper resembles Goerzen et al. (2013) in using Japanese data but they only focus on a few global cities and do not take into account different sectors. Recent papers such as Alfaro and Chen (2014) also raise the issue that spatial boundaries may affect the agglomeration levels in the sense that these boundaries are not continuous like distance. Earlier, Blonigen et al. (2007) also suggests aggregation of country level FDI may affect their results on the US's FDI to European countries. Therefore, we test the robustness our results by using different administrative levels.

Within-firm agglomeration can be explained in the context of the global value chain. With the reduction of trade costs and communication costs, firms can fragment their production into different tasks or functions while at the same time

[1]By group of firms, we mean firms that are linked together through cross ownership.

cities evolve to accommodate certain functions (for example, Grossman and Rossi-Hansberg (2008), Duranton and Puga (2005)). However, face-to-face meetings are still very crucial for production design, marketing, planing and so on (Strauss-Kahn and Vives, 2009).The level of communication depends on the type of affiliates, where headquarters and services will need more face-to-face meetings than production so that they are more likely to be located near each other. While daily face-to-face meetings among the firm's affiliates make production run smoothly, communication among firms within a financial groups could lower transaction costs for the firm in entering a new region. A representative of a foreign exploration and production company in Jakarta stressed "to keep in touch very closely with the authorities here. You need to be physically there. [...] You have to meet people daily from all departments" (Breul, 2019). On the other hand, cities of high connectivity may mitigate the necessity for face-to-face meetings by their advanced producer services (Goerzen et al. (2013), Belderbos et al. (2017)).[2]

Our results show that, in general, within-industry agglomeration in a region significantly increases the possibility of a new affiliate in that region. Group and firm level agglomeration are only significant in certain cases. The first results are quite intuitive and coincide with Defever (2012) and Mayer et al. (2010). The results on group and firm somewhat contradict those of Defever (2012) where the presence of existing affiliates, except for sales, increases the probability of a new affiliate from the same or different sectors. It also contradicts that of Mayer et al. (2010) since their group level agglomeration index always increases the possibility of a new investment. We further test our results by examining different levels of spatial boundaries, and some qualitative conclusions still hold. With regards to city connectivity, we find very strong effects for cities connected by the non-manufacturing members of the firms' financial groups and wholesale companies. However, cities connected by finance, services and transportation companies do not show significant effects in most cases.

The remainder of this paper includes data and descriptive statistics, empirical models, estimation method, results and conclusion.

2 Data and descriptive statistics

We use the Overseas Japanese Companies Data and Domestic Affiliates Database from the Toyo Keizai Database to collect information about Japanese affiliates abroad and groups of Japanese firms. We consider each affiliate that has at least

[2]By connectivity, we mean the interlocking model as in Taylor (2001), where firms connect cities through their activities. Two cities are connected if they have subsidiaries of the same firm, counted as one link. A city with more active firms are more likely to have more links, hence more likely to be well connected.

more than 10 percent ownership from Japanese firm(s) as an investment in a foreign country. We use the year that the affiliate was established as the entry year. (We use the word "having new affiliate", "subsidiary" and "investing" interchangeably.)

To gather information about the spatial locations of affiliates, we use an algorithm to retrieve postal codes from affiliates' addresses to calculate their coordinates. In the case where postal codes are not available, we search for cities in addresses.[3] We start with 4,890 firms and 26,748 affiliates. We end up with 527 firms and 1,233 foreign affiliates after excluding those with less than 10 percent ownership by Japanese firms, those outside the EU and those whose parent companies are not manufacturing firms. These firms are linked with the Domestic Affiliates Database to create information about groups of firms. We are able to identify 49 groups that cover 147 firms (see table 9). We also use the Japan Company Handbook to get information about each firm's capital, year of establishment and employees. We only keep observations of which investments were made after 1990 but keep all possible investments before 1990 to calculate the initial state of each industry, group and firm agglomeration.

To construct city connectivity, we use Japanese firms in four sectors: Finance and Insurance, Service Activities, Transportation and Communications, Wholesale and Retail Trade. To create the connectivity index, we only focus on firms that have more than two or three affiliates in the EU (see table 10)[4]. Specifically, we use 45 finance and insurance firms, 101 services firms, 53 transportation and communication firms, and 119 wholesale and retail firms. We convert sector classifications by the Toyo Keizai Database to SNA sectors (see table 11).

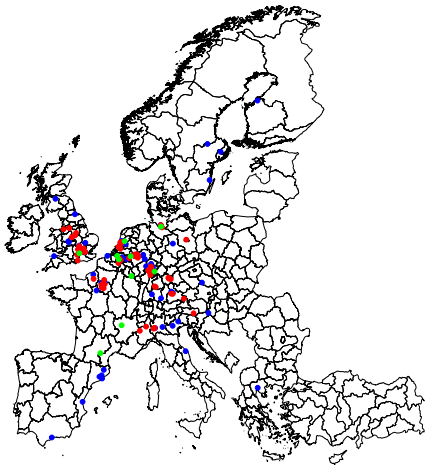
For regional characteristics and region maps, we use data from EuroStat. We use this map to extract information of what regions firms belong to with calculated coordinates. Other data includes population density, land, GDP, value added per employee and human capital as people with tertiary education and/or employed in science and technology as a percentage of total population.

Figure 1 shows locations of affiliates from firms that have more than one affiliate in the EU before and after 1990. Manufacturing and Headquarters seem to be concentrated in certain regions such as the UK, Belgium, Netherlands and some western countries while wholesale affiliates are more scattered over many regions.

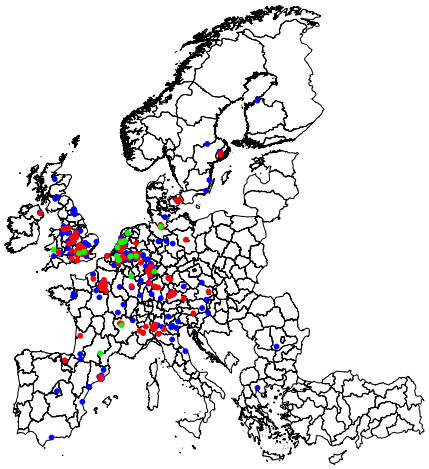
As mentioned earlier, the choice of a location may differ based on region size. Because we have detailed coordinates of affiliates, we can assign different administrative levels including country, group of states, states and group of districts. EuroStat provides statistics for different levels of Nomenclature of Units for Ter-

[3]We use the geonames.org database to search for coordinates from postal codes as well as cities. We also check results manually to make sure they are correct.

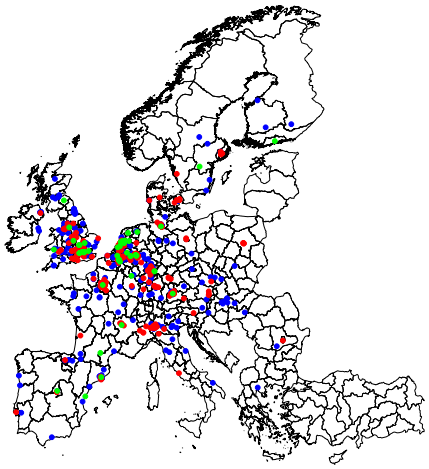
[4]The criteria to choose these firms is the top 25 percentile of the distribution of the number of affiliates of each firm within each sectors. This means we only focus on firms that can create connection among cities.



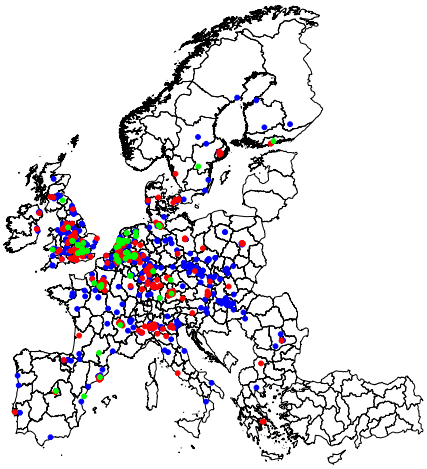
(a) 1980



(b) 1990



(c) 2000



(d) 2010

Figure 1: Accumulated Japanese Affiliates in EU in different years

Manufacture(Blue), Wholesale(Red), Headquarters(Green), Services(Purple)

ritorial Statistics (NUTS), we estimate for four levels from NUTS0 (country) to NUTS3 (group of districts) level. In another estimation, we focus on NUTS2 to examine affiliates' distances from region centers to compare our results with Defever (2012).

3 Empirical models

Following Defever (2012) and Mayer et al. (2010), we assume that firm f choose to open affiliates of sector s in region r at time t if $\pi^s_{frt} > \pi^s_{fr't}$ for $r, r' \in R$ where π^s_{frt} is the expected profit function of the firm. This profit function depends on region r characteristics X_{rt} and four components that reflect within-industry, within-group, within-firm agglomeration effects and connection index.

$$\begin{aligned}\pi^s_{frt} = X_{rt}\boldsymbol{\beta}^s + \delta^s_1 Log\_IndNet^s_{i(f)r(t-1)} + \delta^s_2 Dummy\_IndNet^s_{i(f)r(t-1)} + \\ \gamma^s_1 Log\_GroupNet^s_{fr(t-1)} + \gamma^s_2 Dummy\_GroupNet^s_{fr(t-1)} + \\ \eta^s FirmNet^s_{ir(t-1)} + \\ \sum_S \tau_S SectorConnect^s_{Sr(t-1)} + + \\ \zeta^s GroupConnect_{fr(t-1)} + \epsilon^s_{irt} \qquad (1)\end{aligned}$$

The within-industry indices have two variables: a dummy showing an existing affiliate from the same industry with firm f in region r at time t and a log of accumulated affiliates count plus 1. Separating the dummy and log count allows us to see the effect of the first affiliate and the effect of any additional affiliates of the same industry. We should remind readers that we exclude those affiliates that belong to the same group here. In the baseline specification, we count affiliates of all sectors. To consider the effect of prior affiliates of different sectors on the location choice of a specific sector, we calculate the dummy and log count variables for each separate affiliate sector (we call this the cross-sector effect).

The group level indices also have a dummy showing an existing affiliate belonging to the same financial group as firm i in region r at time t and an index that is calculated as following:

$$Log\_GroupNet^s_{fr(t-1)} = \sum_{g(k), k \neq f} share_{a(g(k)))} D^s_{g(k)r(t-1)} \qquad (2)$$

$D^s_{fr(t-1)}$ is a dummy that equals 1 if firm i has an affiliate of sector s in region r at time $t-1$ and $share_{a(g(k))} = cap_{gk} * cap_{ka}$ is the product of ownership ratio of

group g in firm k with the ownership ratio of firm k in affiliate a. By weighting the ownership ratio instead of just counting the number of affiliates of group g, we take into account the inter-relationship in terms of "ownership control" among different affiliates a that belongs to group g. Similar to within-industry indices, we sum the above index $Log\_GroupNet^{s}_{fr(t-1)}$ over all sectors in the baseline specification and separate each sector to examine the cross-sector effects in a different specification. We exclude affiliates that belong to firm f in both cases.

The firm level index is a dummy indicating the firm's prior investment in region r. To take into account cross-sector effects, we have a list of dummies for different sector s that is equal 1 if there is at least one prior investment in sector s of firm f in region r at time t.

As we have mentioned earlier these three components are proxies for industry, group and firm level agglomeration effects. We expect these coefficients to have positive signs.

The connectivity index consists of two components. The first one is calculated as the log count of affiliates of firms that have at least two affiliates in more than two regions[5]. This first components are further divided into four categories in accordance with four sectors of the firms: Finance and Insurance, Service Activities, Transportation and Communications, Wholesale and Retail Trade. Higher value of the index indicates more services from Japanese firms being available in the region. The second component is the log of ownership-weighted-count of affiliates from non-manufacturing members from the same financial group of the firm[6]. We do not include these firms in the first components to avoid correlation between the two. Similarly, higher value of the index indicates more stronger presence of the firm's group in the region. We expect these indices have positive signs.

X_{rt} includes land (in square kilometers) which is a proxy for resources, population density (inhabitants per square kilometers) which reflects market size, human capital (people with tertiary education and/or employed in science and technology as a percentage of total population) and average value added for employees in industries (one thousand Euros per person). Since we only have firms from one country, we do not include other gravity-like controls such as common language, distance, etc.

[5] "Connectivity" is used loosely in this case. We assume that cities with many affiliates of well connected firms are more likely to be more connected to each other. If we consider the firms and cities in a two-mode network, what we calculate here is the log of in-degree of cities, which is one simple measure of centrality in network analysis (for example, Borgatti and Everett (1997)). In the context of global city as in Taylor (2001), if we consider the value of firm to be 1, then what we calculate here is the log of city nodal connection.

[6] This is different from the Group Level index because these are non-manufacturing firms while the Group Level index just count affiliates of manufacturing firms.

4 Estimation Method

We assume that profits coming from different sectors do not correlate with each other so they can be estimated independently. We further assume that profits are independent among periods. However, there may be unobserved characteristics of firms that make error terms correlated across regions. One way to address this issue is to assume that coefficients $\boldsymbol{\beta}$ are random variables that follow some distribution with mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$. Here, we assume these coefficients follow normal distributions. The model can be estimated using simulated maximum likelihood (Train, 2002). For comparison, we also report results for conditional logit model, however, unobserved firms characteristics, if any, are not taken into account.[7]

To be more specific, we are interested in estimating parameters $(\boldsymbol{\beta}, \boldsymbol{\theta}(\boldsymbol{\gamma}, \boldsymbol{\delta}, \eta, \boldsymbol{\tau}, \zeta))$ given the choice y_{frt} of firm f to region r in time t conditional on region characteristics X_{rt} and industry, group, firm level agglomeration indices and connectivity indices A_{frt}.[8] Assuming error terms follow type I extreme distribution, the choice probability for $y_{rft} = 1$ can be calculated as in equation 3.

$$P_{frt}(y_{frt} = 1) = \int \frac{\exp(X_{rt}\boldsymbol{\beta} + A_{frt}\boldsymbol{\theta})}{\sum_{fqt} \exp(X_{qt}\boldsymbol{\beta} + A_{fqt}\boldsymbol{\theta})} f(\boldsymbol{\beta}|\boldsymbol{\mu}, \boldsymbol{\sigma}) d\boldsymbol{\beta} \tag{3}$$

Here, we assume normal density function $f(\boldsymbol{\beta}|\boldsymbol{\mu}, \boldsymbol{\sigma})$. Since we assume independent choice between periods, we can write the likelihood function taking into account all periods for each firm f as in equation 4.

$$S_f = \int \prod_{t=1}^{T_f} \prod_{r=1}^{R} \left[\frac{\exp(X_{rt}\boldsymbol{\beta} + A_{frt}\boldsymbol{\theta})}{\sum_{fqt} \exp(X_{qt}\boldsymbol{\beta} + A_{fqt}\boldsymbol{\theta})} \right]^{y_{frt}} f(\boldsymbol{\beta}|\boldsymbol{\mu}, \boldsymbol{\sigma}) d\boldsymbol{\beta} \tag{4}$$

Since it is not feasible to integrate out $\boldsymbol{\beta}$ in most cases, we can use the simulation approach described in (Train, 2002) to calculate the simulated log-likelihood and estimate mean $\boldsymbol{\mu}$ and stand deviation $\boldsymbol{\sigma}$ of random parameters $\boldsymbol{\beta}$ and fixed parameter $\boldsymbol{\theta}$. If the estimated variance $\boldsymbol{\sigma}^2$ is tested to be zero, mixed logit will collapse to simple logit.

Because variances of error terms may be different for estimations with different NUTS levels, we should not compare the absolute value of coefficients. We can

[7] Another way to model preference heterogeneity is to allow scale of the idiosyncratic error term to be different among consumers as in Fiebig et al. (2010). These authors develop the so-called generalized multinomial logit model where the random coefficient can be written as $\beta_i = \sigma_i\beta + [\gamma + \sigma_i(1-\gamma)]\eta_i$. σ_i is the individual-specific scale of the idiosyncratic error. We run the baseline regression with this method and find no differences in results so we do not report here.

[8] Sector subscripts are omitted for brevity.

use the ratio among coefficients instead. For example, the ratio δ_2/γ_2 for two dummies of Industry and Firm Level Agglomeration shows how big/small the effects of having prior industry presence on probability of new affiliate compared to prior firm's existing affiliates.

5 Empirical Results

5.1 Basic Specifications

In the basic specification, we estimate four equations for four sectors of affiliates and all agglomeration indices are the sum of all sectors of prior investments. The result is in tables 1, 2 and 3 where each sector has four columns. The first two columns report results without connectivity indices while the last two add these. We also report results from conditional logit and the mixed logit for each specification.

Table 1 shows results for Manufacturing affiliates. The region-specific variables including density, land and value added per employee do not show significant results while Human Capital is significantly positive in all specifications. This suggests that the type of manufacturing is not very skill-intensive. The connectivity indices for four sectors do not show any significance but connectivity among non-manufacturing members of the firm's group is significantly positive. The manufacturing member dummy (Group level dummy) is also significantly positive. This shows that manufacturing subsidiaries tend to be located in regions with strong presence of financial groups, but not necessarily Japanese firms in general. The significantly positive sign of Firm level dummy shows that an existing affiliate of the firm increases the probability of a new investment. But the log count of industry level affiliates with significantly negative sign indicates that costs caused by congestion may have dominated the benefits from agglomeration.

Table 2 shows results for Wholesale affiliates. Density as proxy for market size, Human Capital and Land are all significantly positive as expected. Similarly, Industry level dummy and Firm dummy and Group Connectivity are all significantly positive. However, Group level dummy is no longer significant and Connectivity in Wholesale sector is significantly negative. We might expect firms to spread out their wholesale networks to reach more customers given the fixed costs to open new wholesale affiliate is often not as high as in manufacturing. In many cases, a small representative office would do the job. This finding is similar to Defever (2012).

Table 3 shows results for Headquarters and Services. In the case of Services, since the mixed logit model does not converge, we can only report the conditional logit. Human Capital is significantly positive for Headquarters, which often

requires high-skill workers. Similar to Manufacturing, while having one affiliate from the same industry or group increases the probability of having a new affiliate in the region, more than one industry level affiliate will reduce the probability. Only the Connectivity of the wholesale sector is positive at 5 percent while other sectors such as services, finance or transportation do not show significant results. In the last two columns for Services, only Industry level dummy shows significant results. We have expected a stronger results for these two sectors because Headquarters and Services are more likely to benefit from the connectivity of the region.

In all specifications, the coefficients for agglomeration effects are higher that those of connectivity indices. For example, in table 1, the ratio between Group Connectivity and Indstry level log count is about 0.6. Since the latter is negative, it suggests the effect of having less than one affiliate of the same industry is equal to increasing the Connectivity Index of the region by 0.6 percent.

In the mixed logit models, most standard values for random coefficients are not significant. It may be the case that mixed logit has not captured the unobserved heterogeneity.[9]

[9]We have tried different approaches as in Fiebig et al. (2010) but the results are almost the same, including the insignificance of standard deviation of random parameters. However, the τ in the scale parameter $\sigma_i = \exp \bar{\sigma} + \tau \vee_i$ is significant. This means there are heterogeneous preferences by mean of different error variances of each individual.

Table 1: Location choice at NUTS Level 2

	Manuf.C1	Manuf.M1	Manuf.C2	Manuf.M2
main				
Density	-0.044	-0.050	-0.069	-0.077
	(0.102)	(0.102)	(0.107)	(0.108)
Human Cap.	-0.773**	-0.700*	-0.798**	-0.695*
	(0.271)	(0.293)	(0.276)	(0.301)
Land	0.047	0.044	0.033	0.029
	(0.101)	(0.101)	(0.104)	(0.105)
VA/emp	-0.122	-0.127	-0.097	-0.104
	(0.124)	(0.124)	(0.126)	(0.126)
IndNet(log)	-1.093***	-1.121***	-0.970***	-0.978***
	(0.253)	(0.258)	(0.259)	(0.262)
GroupNet(log)	-0.418	-0.547	0.743	1.020
	(1.690)	(1.686)	(1.620)	(1.687)
IndNet(dum.)	5.360***	5.392***	5.185***	5.204***
	(0.310)	(0.316)	(0.317)	(0.322)
GroupNet(dum.)	2.829**	2.841**	2.173*	2.043*
	(0.967)	(0.974)	(0.944)	(0.966)
FirmNet	7.660***	7.487***	7.571***	7.422***
	(1.100)	(1.114)	(1.110)	(1.125)
SecCon.Fin			0.066	0.057
			(0.146)	(0.147)
SecCon.Ser			-0.066	-0.077
			(0.141)	(0.145)
SecCon.Tra			-0.148	-0.148
			(0.107)	(0.107)
SecCon.Who			-0.053	-0.060
			(0.045)	(0.045)
GroupConnect			0.531**	0.556***
			(0.167)	(0.166)
SD				
Density		0.001		0.001
		(0.171)		(0.171)
Human Cap.		0.501		0.618
		(0.655)		(0.562)
Land		0.001		0.001
		(0.179)		(0.178)
VA/emp		-0.007		-0.006
		(0.238)		(0.234)
Observations	48427	45441	48427	45441

Standard errors in parentheses

C: Conditional Logit, M: Mixed Logit

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Location choice at NUTS Level 2

	Whole.C1	Whole.M1	Whole.C2	Whole.M2
main				
Density	0.612***	0.577***	0.672***	0.653***
	(0.113)	(0.115)	(0.123)	(0.126)
Human Cap.	0.761*	0.705*	0.925**	0.868*
	(0.316)	(0.324)	(0.325)	(0.346)
Land	0.338***	0.326**	0.373***	0.370***
	(0.102)	(0.103)	(0.107)	(0.109)
VA/emp	0.081	0.573*	0.041	0.529
	(0.180)	(0.271)	(0.180)	(0.271)
IndNet(log)	0.072	0.007	0.234	0.168
	(0.176)	(0.186)	(0.181)	(0.192)
GroupNet(log)	4.140	4.460	5.739	5.782
	(3.778)	(4.351)	(4.260)	(4.777)
IndNet(dum.)	4.678***	4.764***	4.540***	4.648***
	(0.256)	(0.274)	(0.262)	(0.280)
GroupNet(dum.)	3.487	3.363	2.369	2.385
	(2.280)	(2.617)	(2.481)	(2.802)
FirmNet	7.448***	7.298***	7.560***	7.501***
	(1.215)	(1.229)	(1.228)	(1.251)
SecCon.Fin			0.040	0.043
			(0.146)	(0.150)
SecCon.Ser			-0.085	-0.078
			(0.103)	(0.104)
SecCon.Tra			-0.129	-0.169*
			(0.081)	(0.083)
SecCon.Who			-0.092**	-0.092**
			(0.031)	(0.032)
GroupConnect			0.448*	0.462*
			(0.222)	(0.229)
SD				
Density		0.045		0.056
		(0.178)		(0.183)
Human Cap.		0.143		0.372
		(1.798)		(1.651)
Land		-0.046		-0.033
		(0.252)		(0.218)
VA/emp		0.958***		0.955***
		(0.206)		(0.205)
Observations	55019	49800	55019	49800

Standard errors in parentheses

C: Conditional Logit, M: Mixed Logit

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Location choice at NUTS Level 2

	Head.C1	Head.M1	Head.C2	Head.M2	Serv.C1	Serv.M1
main						
Density	0.217	0.222	-0.105	-0.106	0.453	0.238
	(0.217)	(0.217)	(0.233)	(0.234)	(0.350)	(0.387)
Human Cap.	2.187**	2.263**	1.498*	1.583*	0.598	0.361
	(0.714)	(0.721)	(0.742)	(0.746)	(1.110)	(1.137)
Land	-0.177	-0.168	-0.412	-0.409	0.502	0.448
	(0.204)	(0.204)	(0.216)	(0.218)	(0.322)	(0.353)
VA/emp	0.463	0.446	0.621	0.591	0.503	0.654
	(0.395)	(0.395)	(0.409)	(0.409)	(0.667)	(0.680)
IndNet(log)	-0.815*	-0.803*	-1.352**	-1.342**	-0.379	-0.988
	(0.389)	(0.391)	(0.467)	(0.470)	(0.513)	(0.615)
GroupNet(log)	-3.359	-3.535	-3.917	-4.117	26.816	30.876
	(3.773)	(3.778)	(3.439)	(3.465)	(14.604)	(18.465)
IndNet(dum.)	5.166***	5.180***	5.436***	5.451***	5.876***	6.571***
	(0.532)	(0.537)	(0.607)	(0.612)	(0.718)	(0.814)
GroupNet(dum.)	3.162	3.210	3.551	3.644	-7.094	-8.020
	(2.405)	(2.411)	(2.312)	(2.319)	(5.567)	(6.803)
FirmNet	21.815	31.865	20.232	23.351	12.686	14.960
	(1230.236)	(183608.933)	(622.388)	(2904.922)	(1070.574)	(2250.846)
SecCon.Fin			-0.694	-0.715		-0.428
			(0.494)	(0.497)		(0.883)
SecCon.Ser			-0.011	-0.009		0.476
			(0.147)	(0.147)		(0.295)
SecCon.Tra			0.039	0.030		-0.231
			(0.127)	(0.129)		(0.336)
SecCon.Who			0.102*	0.103*		0.138
			(0.042)	(0.042)		(0.088)
GroupConnect			0.512	0.520		-0.316
			(0.309)	(0.309)		(0.817)
SD						
Density		-0.003		-0.004		
		(0.271)		(0.316)		
Human Cap.		0.016		0.014		
		(1.774)		(1.556)		
Land		-0.001		-0.001		
		(0.193)		(0.237)		
VA/emp		0.004		0.003		
		(0.511)		(0.519)		
Observations	12080	11929	12080	11929	6795	6795

Standard errors in parentheses

C: Conditional Logit, M: Mixed Logit

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.2 Results at different levels of NUTS

From this section on, we focus on Manufacturing and Wholesale since Headquarters and Services do not always converge. We estimate the same specifications as in previous section but at different levels of NUTS. In the smaller level NUTS3, we do not have data for Human Capital and Value Added per Employee so we just use Density and Land. Results are reported in tables 4 and 5. Each table has seven columns. The first four show results without Connectivity Indices, and the last three add them[10]. The four columns correspond to four administrative levels[11]. In a nutshell, there have been changes in signs of some regional specific characteristics. For our indices, even though magnitudes change, the signs are the same for industry and firm level indices.

Table 4 shows the results for Manufacturing. Human Capital becomes insignificant at country and group of state level. Land becomes significantly positive at group of district level. In terms of significance, agglomeration indices at industry and firm level are almost the same for all specifications except for $GroupNet(log)$. It becomes significantly negative at group of district level while $IndNet(log)$ becomes insignificant at country level. In terms of magnitude, we can look at the ratio between the industry dummy and firm dummy, which are 1.2, 1.3, 1.4 and 1.3 for country, group of states, states, and group of district levels. Regarding connectivity index, $GroupConnect$ becomes insignificant at country and state levels.

Table 5 shows the results for Wholesale. Density, Land and Value Added per employee remain significantly positive. However, as in the case of Manufacturing, Human Capital becomes insignificant at country and group of state level. In terms of significance, industry and firm level indices remain significant. The ratio between firm level and industry level dummy is less than one for country level while others are 1.7, 1.5 and 1.7 consecutively. Group level dummy is only significant at country level but become insignificant for smaller levels. This highlights the results of Mayer et al. (2010) who also studies this effect at country level. Similar to Manufacturing, $GroupConnect$ becomes insignificant at country and state levels.

There may be a few reasons why we observe this feature at Wholesale but not Manufacturing. The latter may need more fixed costs to set up and often cluster with industry while wholesale is less costly to set up and can cover a wide area to reach customers and/or collect market information.

[10]There is no Connectivity for Group of District level because there are collinearity.

[11]The four levels are 0 (country), 1 (Group of States), 2 (States), 3 (Group of Districts).

Table 4: Mixed Logit at Different NUTS levels for Manufacturing

	Manuf.0a	Manuf.0b	Manuf.1a	Manuf.1b	Manuf.2a	Manuf.2b	Manuf.3
Mean							
IndNet(log)	-0.323	-0.260	-0.833***	-0.848***	-1.121***	-0.978***	-1.228**
	(0.190)	(0.197)	(0.211)	(0.218)	(0.258)	(0.262)	(0.411)
GroupNet(log)	1.938	1.970	0.125	0.422	-0.547	1.020	-6.365**
	(1.065)	(1.095)	(1.578)	(1.658)	(1.686)	(1.687)	(2.403)
IndNet(dum.)	4.204***	4.170***	5.017***	5.047***	5.392***	5.204***	5.267***
	(0.328)	(0.332)	(0.290)	(0.296)	(0.316)	(0.322)	(0.415)
GroupNet(dum.)	1.735	1.731	2.080*	1.816*	2.841**	2.043*	6.185***
	(0.897)	(0.931)	(0.877)	(0.926)	(0.974)	(0.966)	(1.129)
FirmNet	5.037***	5.110***	6.671***	6.637***	7.487***	7.422***	6.845***
	(0.648)	(0.682)	(1.113)	(1.118)	(1.114)	(1.125)	(0.884)
Density	-0.005	0.114	-0.102	-0.123	-0.050	-0.077	-0.035
	(0.175)	(0.200)	(0.124)	(0.135)	(0.102)	(0.108)	(0.081)
Human Cap.	0.270	0.525	-0.576	-0.640	-0.700*	-0.695*	
	(0.517)	(0.549)	(0.345)	(0.359)	(0.293)	(0.301)	
Land	0.134	0.167	0.027	0.019	0.044	0.029	0.195*
	(0.093)	(0.099)	(0.110)	(0.115)	(0.101)	(0.105)	(0.081)
VA/emp	-0.400*	-0.450**	-0.108	-0.082	-0.127	-0.104	
	(0.162)	(0.166)	(0.124)	(0.128)	(0.124)	(0.126)	
SecCon.Fin		-0.027		0.013		0.057	
		(0.037)		(0.071)		(0.147)	
SecCon.Ser		0.039		0.034		-0.077	
		(0.036)		(0.070)		(0.145)	
SecCon.Tra		-0.041		-0.109		-0.148	
		(0.035)		(0.061)		(0.107)	
SecCon.Who		-0.018		0.011		-0.060	
		(0.019)		(0.027)		(0.045)	
GroupConnect		-0.043		0.103		0.556***	
		(0.097)		(0.131)		(0.166)	
SD							
Density	0.214	0.279	0.000	0.000	0.001	0.001	0.003
	(0.534)	(0.427)	(0.269)	(0.243)	(0.171)	(0.171)	(0.191)
Human Cap.	1.162	1.243*	0.751	0.734	0.501	0.618	
	(0.628)	(0.610)	(0.655)	(0.678)	(0.655)	(0.562)	
Land	-0.002	-0.001	-0.003	-0.003	0.001	0.001	0.240*
	(0.173)	(0.164)	(0.174)	(0.165)	(0.179)	(0.178)	(0.104)
VA/emp	-0.000	0.000	-0.003	-0.004	-0.007	-0.006	
	(0.226)	(0.224)	(0.210)	(0.220)	(0.238)	(0.234)	
Observations	6852	6852	21828	21828	45441	45441	95148

Standard errors in parentheses

Manuf: Manufacturing, Who: Wholesale. Numbers indicate NUTS level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Mixed Logit at Different NUTS levels for Wholesale

	Whole.0a	Whole.0b	Whole.1a	Whole.1b	Whole.2a	Whole.2b	Whole.3
Mean							
IndNet(log)	-0.191	-0.152	0.188	0.223	0.007	0.168	-0.208
	(0.245)	(0.254)	(0.196)	(0.197)	(0.186)	(0.192)	(0.229)
GroupNet(log)	-0.325	0.356	5.796	5.790	4.460	5.782	5.188
	(1.265)	(1.460)	(3.747)	(3.714)	(4.351)	(4.777)	(3.622)
IndNet(dum.)	4.959***	4.902***	4.475***	4.462***	4.764***	4.648***	4.950***
	(0.455)	(0.458)	(0.297)	(0.300)	(0.274)	(0.280)	(0.281)
GroupNet(dum.)	4.098**	3.685**	1.614	1.619	3.363	2.385	3.138
	(1.315)	(1.358)	(2.114)	(2.082)	(2.617)	(2.802)	(2.129)
FirmNet	3.748***	4.019***	7.804***	7.847***	7.298***	7.501***	8.497***
	(0.677)	(0.759)	(1.361)	(1.374)	(1.229)	(1.251)	(1.106)
Density	0.463*	0.575*	0.505***	0.624***	0.577***	0.653***	0.611***
	(0.221)	(0.248)	(0.143)	(0.168)	(0.115)	(0.126)	(0.079)
Human Cap.	0.863	0.950	0.591	0.798	0.705*	0.868*	
	(0.680)	(0.704)	(0.388)	(0.413)	(0.324)	(0.346)	
Land	0.398**	0.440***	0.430***	0.500***	0.326**	0.370***	0.304***
	(0.121)	(0.126)	(0.126)	(0.137)	(0.103)	(0.109)	(0.071)
VA/emp	0.716*	0.618	0.778**	0.682*	0.573*	0.529	
	(0.343)	(0.350)	(0.274)	(0.279)	(0.271)	(0.271)	
SecCon.Fin		-0.032		0.010		0.043	
		(0.033)		(0.077)		(0.150)	
SecCon.Ser		0.026		-0.021		-0.078	
		(0.034)		(0.066)		(0.104)	
SecCon.Tra		-0.011		-0.035		-0.169*	
		(0.030)		(0.052)		(0.083)	
SecCon.Who		0.003		-0.028		-0.092**	
		(0.020)		(0.025)		(0.032)	
GroupConnect		-0.182		0.017		0.462*	
		(0.149)		(0.179)		(0.229)	
SD							
Density	0.326	0.384	0.228	0.232	0.045	0.056	0.001
	(0.352)	(0.325)	(0.150)	(0.155)	(0.178)	(0.183)	(0.120)
Human Cap.	2.081**	1.861*	0.095	0.200	0.143	0.372	
	(0.802)	(0.835)	(1.247)	(1.597)	(1.798)	(1.651)	
Land	-0.083	-0.090	0.007	0.005	-0.046	-0.033	-0.014
	(0.162)	(0.166)	(0.285)	(0.241)	(0.252)	(0.218)	(0.391)
VA/emp	0.884***	0.840***	0.959***	0.931***	0.958***	0.955***	
	(0.242)	(0.242)	(0.222)	(0.222)	(0.206)	(0.205)	
Observations	7626	7626	24330	24330	49800	49800	106110

Standard errors in parentheses

Manuf: Manufacturing, Who: Wholesale. Numbers indicate NUTS levels.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3 Cross-sector effects

In this section, we replace the indices for all sectors with sector-specific indices to examine the cross sector effects. Results are in table 6. Some of the variables are dropped due to collinearity. The table has eight columns for Manufacturing and Wholesale at four administrative levels. We omit results for Connectivity here since they are almost the same and our focus is agglomeration indices. The first four columns show the results for Manufacturing. In all specifications, Industry dummy and log count and Group dummy of the same sector (ie. Manufacturing) show significant effects while Firm indices show no significance. In terms of signs, dummies are positive but log count shows negative. Wholesale indices have no significant effects on Manufacturing affiliates. The last four columns show results for Wholesale affiliate's location choice. The first three indices of log count are dropped due to collinearity. Industry level indices of both Manufacturing and Wholesale are significantly positive while only the Group level index of Wholesale is significantly positive. The standard errors of firm level indices are very high as well as some other variables. The reason for that may come from some outliers in the data set. However, when we drop the supposed outliers, the cross sector indices also drop due to collinearity.

Nevertheless, we can see that the industry and group level of a sector significantly increase the probability of new affiliates of the same sector at all administrative levels. For cross-sector, wholesale affiliates seem to be located near other manufacturing affiliates but manufacturing affiliates do not tend to be located near other wholesale affiliates. This is somewhat different from Defever (2012) where manufacturing plants tend to be located near other manufacturing and services plants but sales units do not show the same patterns.

Table 6: Cross sector effect of Manufacturing and Wholesale for different NUTS levels

	Manuf.0	Manuf.1	Manuf.2	Manuf.3	Whole.0	Whole.1	Whole.2	Whole.3
choice								
Density	0.331*	0.276*	0.179	0.111	0.384*	0.458***	0.593***	0.640***
	(0.134)	(0.123)	(0.103)	(0.080)	(0.169)	(0.135)	(0.115)	(0.078)
Land	0.185*	0.263*	0.163	0.234**	0.251*	0.305**	0.308**	0.375***
	(0.084)	(0.111)	(0.101)	(0.076)	(0.098)	(0.118)	(0.097)	(0.072)
IndNet(log)Man	-0.552**	-0.924**	-1.526**	-0.487	-0.074			
	(0.208)	(0.343)	(0.558)	(0.906)	(1.725)			
IndNet(log)Who	-47.873	-3.224	-0.055	-5.716	-0.031			
	(66695.431)	(35291.640)	(2976.323)	(23653.125)	(0.166)			
GroupNet(log)Man	0.914	-5.415*	-4.532	-7.402	-1.521			
	(1.264)	(2.338)	(5.640)	(6.591)	(2036875.964)			
GroupNet(log)Who	33.158	-4.041			2.044			
	(126131.598)	(401852.348)			(4.153)			
IndNet(dum.)Man	4.931***	5.363***	6.226***	5.515***	2.082	3.541**	4.040***	5.032***
	(0.335)	(0.374)	(0.529)	(0.812)	(3.064)	(1.112)	(1.105)	(1.223)
IndNet(dum.)Who	32.316	-12.043	-9.709	-6.234	4.773***	4.888***	5.055***	5.116***
	(46229.750)	(28981.152)	(3025.792)	(16859.954)	(0.329)	(0.180)	(0.164)	(0.149)
GroupNet(dum.)Man	1.762*	3.837**	5.237***	7.168***	-17.493	-21.736	-12.178	-12.615
	(0.836)	(1.215)	(1.570)	(1.698)	(1520377.340)	(2799236.436)	(2024.330)	(846.558)
GroupNet(dum.)Who					3.034	6.578***	20.378	21.842
					(2.589)	(1.236)	(719.338)	(1180.068)
FirmNetMan	28.193	21.004	20.345	218.508	-7.488	-1.306	-1.691	-1.385
	(65285.973)	(1708.489)	(961.250)	(34550709.300)	(11490270.479)	(10187.907)	(29806.193)	(25041.374)
FirmNetWho	25.294	53.232	254.473	-8.202	20.213	17.436	19.141	19.904
	(260035.679)	(1.132e+08)	(.)	(790.035)	(3254.224)	(561.958)	(1015.051)	(759.847)
Observations	7212	23052	51432	100602	8242	26434	59028	115499

Standard errors in parentheses

Manuf: Manufacturing, Who: Wholesale. Numbers indicate NUTS level. Blank values indicate being obmitted due to collinearity.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To further examine our results, we follow Defever (2012) to replace the $FirmNet$ dummy with a dummy that equals 1 if there has been an affiliate of the same firm in d km vicinity of the region. However, instead of using geographical centroids, we use the centroid of the most populated cities of that region to calculate the vicinity. In the baseline case, $d = 76km$. To put this in perspective, the minimum distance of an affiliate to a region's centroid is $0.41km$ and the furthest is $244km$. The results are in tables 7 and 8. Again we omit results of Connectivity here to focus on the cross-sector effects. Table 7 shows the results for manufacturing affiliates. At industry level, existing wholesale affiliates do not seems to affect manufacturing location choice. However, wholesale location choice is affected by the existence of both Manufacturing and Wholesale. At group level, existing manufacturing group members' affiliates do increase the likelihood of a new manufacturing affiliate.

The results for firm level are different with Defever (2012) in four points. First, the effects of an existing wholesale unit on setting up a manufacturing subsidiary are only significantly positive at short distances, specifically shorter than $51km$ (In his analysis the effects are still significant up to $120km$). Second, the effect of an existing manufacturing plant on setting up a manufacturing subsidiary has two peaks at $51km$ and $126km$ (In his analysis, the effects die down until $360km$ and rise to $530km$ before dying down again.). Third, the effects of existing wholesale on opening a new wholesale subsidiary are significantly negative for longer distances (In his analysis the coefficients are mostly insignificant.). Fourth, an existing manufacturing affiliate has a positive effect on wholesale only at $76km$ away.

There are several reasons why our results are different from Defever (2012). First, we only use Japanese firms while they use firms from many countries including Japan. In fact, country samples seem to affect the estimation of FDI in many cases (for example, see Blonigen et al. (2007)). Furthermore, it is possible that we have not been able to address properly the source of heterogeneity in firms' preferences[12] . Second, our model controls for sector-specific industry-level and group-level effects that are not present in his model. It is possible that his cross-sector dummy absorbs somewhat this type of effects because firms may want to locate near their group's affiliates too. Indeed, in table 6, we can see that the dummy for cross-sector industry and group index are significantly positive for all administrative levels in the case of manufacturing location choice.

[12] A fixed effect mixed logit may perhaps address the problem.

Table 7: Cross sector effect of prior presence in vicinity for Manufacturing Location

	26km	51km	76km	101km	126km	151km	176km
choice							
Density	0.228*	0.228*	0.231*	0.234*	0.240*	0.237*	0.236*
	(0.104)	(0.104)	(0.104)	(0.105)	(0.105)	(0.105)	(0.104)
Human Cap.	-0.339	-0.319	-0.321	-0.316	-0.311	-0.323	-0.331
	(0.271)	(0.271)	(0.271)	(0.271)	(0.272)	(0.271)	(0.271)
Land	0.203	0.207	0.211*	0.215*	0.223*	0.219*	0.213*
	(0.106)	(0.106)	(0.106)	(0.106)	(0.106)	(0.106)	(0.106)
VA/emp	-0.076	-0.083	-0.081	-0.084	-0.088	-0.083	-0.076
	(0.120)	(0.120)	(0.120)	(0.120)	(0.120)	(0.120)	(0.120)
IndNet(log)Man	-1.881**	-1.890**	-1.867**	-1.878**	-1.900**	-1.861**	-1.824**
	(0.575)	(0.576)	(0.575)	(0.576)	(0.578)	(0.576)	(0.573)
IndNet(log)Who	0.153	0.155	0.173	0.170	0.165	-0.151	-0.047
	(1966.673)	(2208.308)	(1698.640)	(3180.200)	(2183.337)	(1454.987)	(1521.899)
GroupNet(log)Man	-4.970	-4.969	-4.973	-4.981	-5.038	-5.017	-4.974
	(5.705)	(5.700)	(5.696)	(5.695)	(5.680)	(5.680)	(5.685)
IndNet(dum.)Man	6.536***	6.537***	6.519***	6.528***	6.533***	6.507***	6.479***
	(0.539)	(0.540)	(0.539)	(0.540)	(0.542)	(0.540)	(0.538)
IndNet(dum.)Who	-8.966	-9.223	-8.750	-9.995	-9.238	-8.486	-8.572
	(2039.937)	(2287.023)	(1752.852)	(3279.630)	(2250.177)	(1591.007)	(1630.802)
GroupNet(dum.)Man	5.295***	5.304***	5.305***	5.316***	5.370***	5.345***	5.309***
	(1.562)	(1.561)	(1.561)	(1.561)	(1.558)	(1.560)	(1.562)
Man in 26km	0.905						
	(0.727)						
Who in 26km	3.472**						
	(1.158)						
Man in 51km		1.085*					
		(0.512)					
Who in 51km		2.829*					
		(1.108)					
Man in 76km			0.652				
			(0.493)				
Who in 76km			2.030				
			(1.059)				
Man in 101km				0.720			
				(0.421)			
Who in 101km				1.719			
				(1.060)			
Man in 126km					0.910**		
					(0.347)		
Who in 126km					1.464		
					(1.066)		
Man in 151km						0.641	
						(0.343)	
Who in 151km						1.207	
						(1.062)	
Man in 176km							0.208
							(0.342)
Who in 176km							0.867
							(1.060)
Observations	48427	48427	48427	48427	48427	48427	48427

Standard errors in parentheses

Man: Manufacturing, Who: Wholesale.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Cross sector effect of prior presence in vicinity for Wholesale Location

	26km	51km	76km	101km	126km	151km	176km
choice							
Density	0.588***	0.588***	0.586***	0.585***	0.580***	0.576***	0.576***
	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)	(0.116)
Human Cap.	0.831**	0.828**	0.841**	0.834**	0.826**	0.818**	0.814**
	(0.307)	(0.307)	(0.307)	(0.307)	(0.307)	(0.307)	(0.307)
Land	0.312**	0.305**	0.305**	0.302**	0.295**	0.289**	0.279**
	(0.106)	(0.107)	(0.107)	(0.107)	(0.107)	(0.107)	(0.107)
VA/emp	0.106	0.112	0.109	0.113	0.119	0.127	0.135
	(0.169)	(0.169)	(0.169)	(0.169)	(0.169)	(0.169)	(0.169)
IndNet(dum.)Man	3.960***	3.967***	3.976***	3.968***	3.959***	3.886***	3.888***
	(1.103)	(1.103)	(1.103)	(1.103)	(1.104)	(1.106)	(1.106)
IndNet(dum.)Who	4.980***	5.005***	5.002***	5.004***	5.010***	5.019***	5.035***
	(0.174)	(0.175)	(0.174)	(0.175)	(0.175)	(0.175)	(0.176)
GroupNet(dum.)Man	-11.886	-11.634	-11.686	-11.678	-11.672	-11.624	-11.621
	(2654.165)	(2363.099)	(2455.512)	(2457.418)	(2459.421)	(2494.262)	(2504.175)
GroupNet(dum.)Who	19.606	19.300	19.310	19.300	19.283	19.267	19.226
	(622.615)	(531.068)	(534.857)	(531.804)	(526.105)	(520.526)	(508.155)
Man in 26km	-11.740						
	(1817.164)						
Who in 26km	-0.277						
	(0.466)						
Man in 51km		-11.476					
		(1203.112)					
Who in 51km		-0.773					
		(0.416)					
Man in 76km			2.337*				
			(1.166)				
Who in 76km			-0.706				
			(0.393)				
Man in 101km				1.868			
				(1.200)			
Who in 101km				-0.828*			
				(0.388)			
Man in 126km					1.664		
					(1.212)		
Who in 126km					-0.975*		
					(0.385)		
Man in 151km						1.350	
						(1.207)	
Who in 151km						-1.068**	
						(0.366)	
Man in 176km							1.077
							(1.212)
Who in 176km							-1.045**
							(0.327)
Observations	55019	55019	55019	55019	55019	55019	55019

Standard errors in parentheses

Man: Manufacturing, Who: Wholesale.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Conclusion

We have examined the effects of within-firm agglomeration and cross-city connectivity on the probability of having a new affiliate in the region. Using data about Japanese firms investing in the EU, we confirm the within industry agglomeration force described in the literature. Furthermore, we separate the effect of affiliates of group companies from that of the firm only. It is possible that having an affiliate of group companies in a region that is far from the firm's prior investment will increase the possibility of the firm investing in that region. In fact, we find both group and firm level indices increase the possibility. In terms of magnitude, the effect of the firm-level index is higher than that of the group-level index for Manufacturing and Wholesale.

We also examine how the role of cities in a "city network" formed by interaction among Japanese firms in different sectors as well as non-manufacturing members of firms' financial group. Cities with more existing non-manufacturing members are more likely to attract wholesale and manufacturing through the capacity to provide more advanced services to the firm's new affiliate.

To take into account the concerns of agglomeration effects being sensitive to geographical boundaries, we have estimated the model at country as well as group of districts levels. In most cases, we do not see any qualitative changes in industry level and firm level dummy coefficients but group level coefficients become insignificant at country and group of state levels. The Connectivity through Group of firms is also insignificant at these levels.

Finally, we have examined the presence of cross-sector affiliates on the probability of having a new affiliate from a separate sector. Here, our results depart from previous literature where we do not see the significance of such effects. This could be because underlying firm distribution of our data is different from others'. We will address this issue in our future research.

References

Alfaro, L. and Chen, M. X. (2014). The global agglomeration of multinational firms. *Journal of International Economics*, 94:263–276.

Antras, P. and Yeaple, S. R. (2015). *Multinational Firms and the Structure of International Trade*, volume 4. Elsevier B.V.

Belderbos, R., Du, H. S., and Goerzen, A. (2015). Global city connectivity and the establishment of MNC regional headquarters. *Conference on Location Decisions of Multinational Enterprises: Market, Cities or Clusters?*, pages 1–39.

Belderbos, R., Du, H. S., and Goerzen, A. (2017). Global Cities, Connectivity, and the Location Choice of MNC Regional Headquarters. *Journal of Management Studies*, 54(8):1271–1302.

Blonigen, B. A., Davies, R. B., Waddell, G. R., and Naughton, H. T. (2007). FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review*, 51:1303–1325.

Borgatti, S. P. and Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks*, 19(3):243–269.

Breul, M. (2019). Cities in 'multiple globalizations': insights from the upstream oil and gas World City Network. *Regional Studies, Regional Science*, 6(1):25–31.

Defever, F. (2012). The Spatial Organization of Multinational Firms. *Canadian Journal of Economics*, 45(2).

Duranton, G. and Puga, D. (2005). From sectoral to functional urban specialisation. *Journal of Urban Economics*, 57:343–370.

Fiebig, D. G., Keane, M. P., Louviere, J., and Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science*, 29(3):393–421.

Goerzen, A., Asmussen, C. G., and Nielsen, B. B. (2013). Global cities and multinational enterprise location strategy. *Journal of International Business Studies*, 44(5):427–450.

Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading Tasks: A Simple Theory of Offshoring. *American Economic Review*, 98(5):1978–1997.

Head, K. and Mayer, T. (2004). Market Potential and the Location of Japanese Investment in the European Union. *Review of Economics and Statistics*, 86(4):959–972.

Head, K., Ries, J., and Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics*, 38(3-4):223–247.

Mayer, T., Mejean, I., and Nefussi, B. (2010). The location of domestic and foreign production affiliates by French multinational firms. *Journal of Urban Economics*, 68(2):115–128.

Strauss-Kahn, V. and Vives, X. (2009). Why and where do headquarters move? *Regional Science and Urban Economics*, 39(2):168–186.

Taylor, P. J. (2001). Specification of the World City Network. *Geographical Analysis*, 33(2):181–194.

Train, K. E. (2002). *Discrete choice methods with simulation*. Cambridge University Press.

Appendix

Table 9: Data cleaning

	dataRange	firmNo	affiNo
1	all	4966	29083
2	Europe	1119	3212
3	FDI	1035	1862
4	with year	1035	1862
5	with geocode	1000	1763
6	parent manuf only	688	1233
7	invest before 1990 only	161	223
8	invest before and after 1990	143	604
9	invest after 1990 only	384	605

Table 10: Numbers of alliliates of each non-manufacturing big firm in the EU

	firmsnasector	Nfirm	min	max	Q25	Q50	Q75
1	Agriculture, Forestry and Fishing	5	1	6	1.00	2	4
2	Construction	14	1	6	1.00	1	2
3	Electricity, gas and water supply	10	1	7	1.25	3	3
4	Finance and insurance	45	1	13	1.00	2	3
5	Manufacturing	761	1	46	1.00	1	3
6	Mining	6	1	2	1.25	2	2
7	Real estate	5	1	3	1.00	1	2
8	Service activities	101	1	15	1.00	1	2
9	Transport and communications	53	1	22	1.00	1	3
10	Wholesale and retail trade	119	1	69	1.00	1	2

Table 11: List of Industries and Sectors

Manufacturing	Wholesale
Chemistry	Chemical wholesale
Drug	Electric equipment wholesale
Electric equipment	Fiber clothes wholesale
Fiber clothes	General Wholesale
Glass soil and stone	Glass soil and stone wholesale
Grocery	Grocery wholesale
Iron and steel	Machinery wholesale
Machine	Other wholesale
Metal products	Petroleum fuel wholesale
Non-ferrous metal	Pharmaceutical wholesale
Other manufacturing industry	Precision equipment wholesale
Petroleum coal	Steel & Metal Wholesale
Precision mechanical equipment	Transportation equipment wholesale
Pulp paper	
Rubber product	
Transport equipment	

Services	Finance
Advertisement	Bank
Architectural Design	Commodity futures
Building management security	Investment Management
Communication broadcasting	Investment Services , etc.
Consulting	Lease
Hotel	Life insurance
Information system software	Money Lending Credit Card
Leisure entertainment	Other financial
Machinery repair	Property and casualty insurance
Newspaper publishing	Securities
Other services	Trust bank
Real estate	
Temporary staffing business contract	
Travel	
Video and Music	

Note: These industries are classified by Toyo Keizai Database. Sectors are in concordance with SNA sectors except for Wholesale because SNA does not separate Wholesale and Retail. Headquarters only contain Headquarters.