Grant-in-Aid for Scientific Research (S)

Real Estate Markets, Financial Crisis, and Economic Growth

: An Integrated Economic Approach

Working Paper Series No.46

The Localization of Interrm Transaction Relationships and Industry Agglomeration

Kentaro Nakajima
Yukiko Umeno Saito
and
Iichiro Uesugi

March, 2016

HIT-REFINED PROJECT
Institute of Economic Research, Hitotsubashi University
Naka 2-1, Kunitachi-city, Tokyo 186-8603, JAPAN
Tel: +81-42-580-9145
E-mail: hit-refined-sec@ier.hit-u.ac.jp
http://www.ier.hit-u.ac.jp/hit-refined/

The Localization of Interfirm Transaction Relationships and Industry Agglomeration*

Kentaro Nakajima† Yukiko Umeno Saito‡ Iichiro Uesugi§

March 12, 2016

Abstract

Using a unique and massive dataset on firms' suppliers and customers, we examine the localization of transaction relationships to find the following. First, based on a counterfactual that controls for the location of firms and their potential partners, transaction relationships in about 90 to 95% of the three-digit manufacturing industries are localized within 40km. Second, based on a counterfactual that controls for the average distance of transaction relationships in the entire manufacturing sector, in about 40% of industries transaction relationships are localized. Third, the extent of industry agglomeration and the extent of the localization of transaction relationships are positively correlated.

Keywords: Interfirm transactions; agglomeration; transaction distance
JEL classification: R11

*We would like to thank Ryo Nakajima, Tsutomu Watanabe, Ralph Paprzycki, and members of the RIETI study group on financial and industry networks for helpful comments. We gratefully acknowledge financial support from the Japan Society for the Promotion of Science (#22730191, #22330073, #23730228, #25380275, #25220502, #26285051, #15H03344) and a grant from the Japan Center for Economic Research.

†Graduate School of Economics and Management, Tohoku University.
‡Research Institute of Economy, Trade, and Industry
§Institute of Economic Research, Hitotsubashi University.

1 Introduction

It is a widely observed fact that there is a strong tendency for industrial activities to be localized in certain areas. A famous example is the agglomerations of automobile assemblers and suppliers in places such as Toyota City in Japan and Detroit in the United States. In fact, looking at empirical evidence for Britain and Japan respectively, Duranton and Overman (2005) and Nakajima, Saito, and Uesugi (2010a) show that about half of all manufacturing industries tend to be localized. A comprehensive survey on the micro-foundations of agglomeration by Duranton and Puga (2004) suggests that interfirm transactions represent one of the most important reasons for industry localization. A number of empirical studies using industry-level data have further highlighted the importance of such transactions. Employing input-output tables for the United States, Rosenthal and Strange (2001), for instance, show that stronger transaction relationships within particular industries contribute to localization in that industry. Ellison, Glaeser, and Kerr (2010), using the same input-output tables, find a positive association between the extent of co-agglomeration of two different industries and the amount of transactions between these industries. These studies imply that a reduction in interfirm transaction costs is one of the major motivations for firms to locate close to each other.

However, mainly due to the lack of data, these previous studies fail to provide any evidence that locating in clusters does indeed reduce firms' transaction costs. However close firms are located to each other, their transaction costs do not necessarily decrease unless they actually have relationships with proximate transaction partners. Further, the use of aggregated data in existing studies due to the lack of more detailed data on inter- or intra-industry transactions may mask the true causality between transaction relationships and industry agglomeration. It is possible that the number of transaction relationships between (within) industries (extensive margin) has a greater influence on industry agglomeration than the value of the transactions within these relationships (intensive margin). However, data from input-output tables contain no information on the extensive margin of these transaction relationships and thus do not allow us to examine

which of the two—the extensive or the intensive margin—are of greater importance in determining industry agglomeration.

Using a unique dataset of 142,282 manufacturing firms in Japan containing information on firms' suppliers and customers from all industries, that is, suppliers and customers hailing from both manufacturing and non-manufacturing industries, we provide the first comprehensive examination of the physical distances between firms engaged in interfirm transaction relationships and discuss how such relationships are localized. However, to do so, simply measuring the physical distances of interfirm relationships is not enough. For example, Figure 1 shows how interfirm transaction relationships are distributed in Japan. The mean distance between manufacturing firms and their transaction counterparts, i.e., suppliers and customers, is 153.3km. The distances for the 25, 50, and 75 percentile are 8km, 39km, and 246km, respectively.

Figure 1

But on their own, these figures do not mean very much. Instead, we need to evaluate if the distances are short or long, in other words, we need to examine whether these transaction relationships are localized. For this purpose, it is necessary to employ an appropriate yardstick against which to compare the actual distribution of interfirm transaction distances, which, in our case, is provided by Duranton and Overman's (2005) pairwise distance approach used for detecting geographical localization of firms. Duranton and Overman's method consists of two steps. First, the distribution of bilateral distances of firms in an industry is drawn. Second, the distribution is then compared with a counterfactual in which firms randomly choose their location from among all manufacturing industry sites, and the statistical significance of the departure from randomness is tested.

In this paper, we focus on bilateral transaction relationships between firms and draw the distribution of these relationship distances. Since our interest lies in interfirm transactions, we focus on the distances of these transaction relationships rather than on the distances between all the pairs of firms in an industry. Further, in order to examine the localization of transaction relationships in two different ways, we employ two types of

counterfactuals and compare them with the distribution of transaction relationship distances: a counterfactual that focuses on potential transaction relationships conditional on the location of firms (location-based counterfactual) and a counterfactual that solely focuses on the actual transaction relationship in the manufacturing sector as a whole (relationship-based counterfactual).

More specifically, the location-based counterfactual uses the information on the location of firms in an industry and of all the firms which can potentially transact with them. The distribution of distances between firms and their potential partners, which we define for each industry, represents the tendency for potential transaction relationships to localize. Using the counterfactual, we investigate the relevance of geographical proximity in determining transaction relationships for each industry. If distance matters, we expect to observe that firms in the industry tend to choose relationships with firms located nearby from the pool of potential transaction partners. In contrast, the relationship-based counterfactual closely follows the spirit of Duranton and Overman (2005) and represents the overall tendency for transaction relationships to localize in manufacturing industry as a whole. Using the counterfactual, we measure the extent of departure from the overall tendency of transaction relationships to agglomerate in manufacturing industry as a whole. Since there is, as we will show, considerable heterogeneity in transaction relationships, measuring the departure from the manufacturing industry average is one way to detect such heterogeneity regarding the localization of transaction relationships. Our main findings are as follows. First, using the location-based counterfactual, we find that in about 90 to 95% of the 150 three-digit industries transaction relationships are localized at a distance of 40km or less, indicating that in most industries physical proximity is a very important factor in firms' choice of transaction partners. The relevance of proximity in interfirm transaction relationships is further highlighted when we compare our results with those of Murata, Nakajima, Okamoto, and Tamura (2011) focusing on patent citation relationships in the United States. Applying Duranton and Overman's approach in a similar manner to ours, they find that in more than 95% of technology categories these relationships are localized at least once within a distance of about 1,200km. While the two

studies are not directly comparable since they focus on different countries, the contrast in distances of localization—40km for interfirm transaction relationships and 1,200km for patent citations—suggest that infirm transactions are more closely associated with industrial agglomeration than knowledge spillovers in the form of patent citations.

Second, using the relationship-based counterfactual, we find that in about 40% of industries relationships are localized at a distance of 40km or less. The transaction relationships of firms in these industries are more concentrated at short distances than those of firms in manufacturing industry as a whole. And third, using the relationship-based counterfactual, we find that there is a positive correlation between the extent of industry localization and the localization of transaction relationships. However, transaction relationships are not necessarily localized in industries that are geographically localized and vice versa. In some of the industries that we identified as being agglomerated the distribution of transaction relationship distances was skewed toward the right relative to the industry average, while in other industries, where transaction relationships were localized at short distances show no notable agglomeration tendency.

The remainder of the paper is organized as follows. Sections 2 and 3 describe our firm-level dataset and empirical approach, respectively. Section 4 then provides the empirical results, while Section 5 concludes.

2 Data

The dataset we use is compiled by a major credit research firm, Tokyo Shoko Research Incorporated (TSR). The dataset includes 826,169 large and small corporations in Japan and consists of two subsets: a dataset on firms' characteristics and a dataset on interfirm relationships. Necessary information for the dataset is collected by field researchers of TSR, who not only utilize public sources such as financial statements, corporate registrations, and public relations documents, but also implement face-to-face interviews with firms, their customers and suppliers, and banks which extend loans to them.

The sub-dataset on firm characteristics includes information on a firm's name, address,

industry classification code,[1] products, year of establishment, number of employees, sales, business profit, and credit score. The other sub-dataset on interfirm relationships includes information on the names of suppliers and customers of a firm.[2] There exists an upper limit of 24 with regard to the number of counterparts each firm can report as its customers or suppliers. The total number of interfirm relationships is approximately four million.

This dataset has several unique features. First, it covers about half of the total of 1.52 million incorporated firms[3] in Japan. Since each of these roughly 830,000 firms reports the names of its customers and suppliers, this dataset makes it possible to describe actual interfirm relationships in all industries in Japan more comprehensively than with any other dataset before. Note, however, that not all transaction relationships are covered in the dataset because of the upper limit on the number of transaction counterparts each firm can report.

Second, by combining the two sub-datasets on firm characteristics and interfirm relationships, we have information on the characteristics of the customers and suppliers of each firm. Furthermore, the dataset includes information on firms' location, which enables us to calculate the distance between two firms engaged in a transaction relationship. In order to identify the geographical location of each firm, we geocode firms' address data using the CSV Address Matching Service provided by the Center for Spatial Information Science, University of Tokyo.[4]

To examine the localization of transaction relationships, we follow previous studies on industry localization and concentrate on the manufacturing sector only, which reduces the number of firms in our dataset to 142,282. While the sample is limited to manufacturing firms and the number of observations we use for analysis is 142,282 throughout the paper, the transaction partners of these manufacturing firms do not necessarily all hail from the manufacturing sector. In fact, many belong to other industries such as wholesale and services. Therefore, we employ data on the transaction relationships between firms in

[1]Industry classifications follow the Japanese Standard Industry Classification (JSIC).

[2]The dataset also has information on the names of major shareholders of a firm. However, in this paper we only focus on transaction relationships and do not use the information on shareholders.

[3]Statistics Bureau, 2004 Establishment and Enterprise Census of Japan.

[4]http://newspat.csis.u-tokyo.ac.jp/geocode

the manufacturing sector and their transaction counterparts, which may well be firms in non-manufacturing industries.

3 Empirical Approach

This section provides an overview of our empirical approach, which closely follows Duranton and Overman's (2005) point-distance method. However, while they measure the industry localization of manufacturing firms, we measure the localization of their transaction relationships. Our kernel density approach á la Duranton and Overman (2005) consists of three steps. First, we calculate the pairwise distances between a firm in a particular industry and its transaction partners. These transaction partners can be either manufacturing or non-manufacturing firms. We then estimate a kernel density function of the distance distribution. Second, in order to implement statistical tests, we construct two types of counterfactuals. The first counterfactual uses the location information of firms and their potential partners and calculates the distances between them. The counterfactual randomly chooses from the pool of such potential transaction relationships. As previous studies (e.g., Duranton and Overman, 2005; Nakajima, Saito, and Uesugi, 2010a) have shown, the geographical distribution of firms or establishments itself is localized. Thus, by using this location-based counterfactual, we control for the fact that there is a tendency for firms to agglomerate. The second counterfactual we employ randomly chooses from the pool of actual transactions between firms and their transaction partners. In other words, we use the overall tendency of transaction relationships to be localized as a benchmark. Third, based on these two counterfactual distance distributions, we construct two confidence interval bands and test whether the transaction relationships in an industry can be considered to be localized.

3.1 Kernel Densities

We begin by estimating the density distribution of pairwise distances between transaction partners. Let S^A be the set of firms in industry A, and n_A be the number of elements,

which in our case is the number of firms in the industry. The set of transaction partners of firm i in industry A is denoted by S_i, and the number of these partners is denoted by n_i. It is worth noting that firms in set S_i can fall into either the manufacturing or the non-manufacturing sector. The great circle distance between firm i and its transaction partner j is denoted by d_{ij}. We then estimate the kernel-smoothed densities (K-densities) of the pairwise distances between transaction partners. The estimator of the K-density at distance d is

$$\hat{K}_A(d) = \frac{1}{h \sum_{i \in S^A} n_i} \sum_{i \in S^A} \sum_{k \in S_i} f\left(\frac{d - d_{ik}}{h}\right), \tag{1}$$

where h is the bandwidth and f is the kernel function.[5]

3.2 Counterfactuals

In this subsection, we construct the counterfactuals used to test the localization of transaction relationships. As mentioned, we employ two types of counterfactuals: a location-based counterfactual and a relationship-based counterfactual.

3.2.1 Location-based Counterfactual

We start with what we call the location-based counterfactual. To construct this counterfactual, we consider the location of each firm in industry A, define potential transaction partners for it, and then calculate the distances between each firm and its potential transaction partners. In this way, we control for the tendency of firms and their potential transaction partners to geographically localize. If all firms and their potential partners are located close to each other, the counterfactual distribution of relationship distances is skewed toward the short end of distances. Thus, the test of localization using the location-based counterfactual focuses on the departure from the randomness in which firms in industry A choose from the pool of potential transaction partners. This test is useful for examining the importance of physical distances in determining transaction relationships in each industry after controlling for the geographical localization of firms

[5]Following Silverman (1986), we use a Gaussian kernel with optimal bandwidth.

in the industry and their potential transaction partners.

Let us explain the procedure in more detail. A firm in industry A transacts not only with firms in its own industry but also in other industries. For each firm, we choose its potential transaction partners from the pool of firms of the industry that the actual partners belong to. If firm i in industry A has an actual transaction relationship with a firm in industry B, the firm's potential transaction partner is randomly chosen from the pool of firms in industry B. For each of the n_i transaction partners for firm i, we randomly choose a potential partner. We then calculate the n_i relationship distances between the firm and its potential transaction partners. We calculate such distances for every firm $i \in S^A$ and estimate a counterfactual K-density. Then we repeat the above procedure 1,000 times and construct confidence intervals.

3.2.2 Relationship-based Counterfactual

Next, let us describe the relationship-based counterfactual, which solely focuses on transaction relationships but not on firms' location. The counterfactual is similar to the one considered by Duranton and Overman (2005) in that it considers the tendency of manufacturing industry as a whole to localize. This counterfactual therefore can be used to examine the departure from the tendency of transaction relationships in manufacturing industry as a whole to be localized, which is useful for detecting inter-industry heterogeneity in the localization of transaction relationships.

We start the construction of the counterfactual by pooling all the transaction relationships of all manufacturing firms. We then randomly choose n_i transaction relationships for firm i from the pool of these actual transaction relationships. In other words, the counterfactual is based on the assumption that firms in a particular industry, say industry A, choose their transaction relationships subject to the average tendency of transaction patterns in the manufacturing sector as a whole. Picking up such potential transaction relationships for every firm $i \in S^A$, we can estimate the counterfactual K-density in industry A. We repeat the above procedure 1,000 times in order to have 1,000 counterfactual K-densities, which are used for constructing confidence intervals.

3.3 The Localization of Transaction Relationships

To statistically test the localization of transaction relationships, we construct two-sided confidence intervals containing 95% of the randomly drawn K-densities. Following Duranton and Overman (2005), we employ local and global confidence bands. Local confidence bands are obtained by selecting the 5th and 95th percentiles of the simulated 1,000 counterfactual K-densities at each distance d, which are labeled as the upper confidence band $\overline{K}_A(d)$ and the lower confidence band $\underline{K}_A(d)$. The interval between $\overline{K}_A(d)$ and $\underline{K}_A(d)$ is the 95% local confidence interval band of industry A at distance d.

Since these local confidence bands only provide statements at each distance d, we calculate global confidence bands, which we use to measure the deviation of K-densities from the counterfactuals over the entire range of distances, which in our case is 0 to 180 km.[6] The global confidence bands are defined so that 95% of the 1,000 randomly drawn K-densities lie above the lower band and another 95% of the randomly drawn K-densities lie below the upper band. Hence, we obtain the upper global confidence band $\overline{\overline{K}}_A(d)$ and the lower global confidence band $\underline{\underline{K}}_A(d)$ for industry A. If $\hat{K}_A(d) > \overline{\overline{K}}_A(d)$ for at least one $d \in [0, 180]$, we can say with 95% confidence that transactions in industry A are globally localized. On the other hand, if $\hat{K}_A(d) < \underline{\underline{K}}_A(d)$ for at least one $d \in [0, 180]$, and industry A is not defined as localized, transaction relations in industry A can be considered to be globally dispersed.

In addition to examining whether a specific industry is localized or dispersed, we also measure the extent of localization or dispersion for each industry at each distance. Specifically, and again following Duranton and Overman (2005), we define the following index of transaction relationship localization at each distance d:

$$\Gamma_A(d) \equiv \max\left(\hat{K}_A(d) - \overline{\overline{K}}_A(d), 0\right), \tag{2}$$

In addition, we define the following index of transaction relationship dispersion for each

[6] We follow Nakajima, Saito, and Uesugi (2010a) in setting the upper bound at 180km in order to obtain results on the extent of localization among manufacturing industries in Japan.

industry at each distance:

$$\Psi_A(d) \equiv \begin{cases} \max\left(\underline{\underline{K}}_A(d) - \hat{K}_A(d), 0\right) & \text{if } \sum_{d=0}^{d=180} \Gamma_A(d) = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

We further define the indices of the extent of transaction relationship localization and dispersion for each industry as Γ_A and Ψ_A, respectively. We do this by summing up $\Gamma_A(d)$ and $\Psi_A(d)$ for all values of $d \in [0, 180]$.

4 Results

4.1 Summary Statistics

Let us start our examination by looking at summary statistics for the transaction relationships of the 142,282 firms in our dataset. As shown in the introduction, the median relationship distance is 39km, while the mean distance is much larger at 153.3km. This is consistent with the shape of the distribution of transaction relationship distances, which is shown in Figure 1. The distribution is skewed toward the shortest end of the scale, but there are also a small but non-negligible number of observations at the longer end. **Figure 2(a) shows the mean, and Figure 2(b) shows 25% percentile, median, and 75% percentile values of transaction relationship distances for three-digit manufacturing industries. In Figure 2(b), solid line represents median, lower dashed line represents 25 percentile, and upper dashed line represents 75 percentile.** There appears to be considerable inter-industry heterogeneity in terms of the distances of actual transaction relationships. Let us therefore have a closer look at the ten three-digit manufacturing industries with the smallest median transaction relationship distances (Table 1) and the largest ones (Table 2). Doing so indicates that the majority of (three-digit) industries in the two tables hail from a small number of broader (two-digit) industries. Specifically, four of the industries with the shortest median transaction distances fall under the heading of printing businesses (JSIC16) and a further two under

leather processing industries (JSIC21). On the other hand, of the industries with the largest median transaction distances, seven belong to the textile and apparel industries (JSIC11 and 12).

Figures 2(a) and 2(b) and Tables 1 and 2

Next, while Figures 2(a) and 2(b) showed actual transaction distances, Figures 3(a) and 3(b) shows summary statistics of potential transaction distances. For the manufacturing sector as a whole, the median relationship distance is 460 km and the mean is 543.5 km, both of which are much larger than the actual distances. Another difference from the actual relationship distances is the smaller gap between the mean and the median, indicating that the distribution of potential relationship distances is less skewed to the left than that of actual relationship distances.

Figures 3(a) and 3(b)

4.2 Localization of Transaction Relationships Relative to Location-based Counterfactual

This and the next subsection present our empirical results based on the approach introduced in Section 3. We estimate the K-density distribution of interfirm transaction relationship distances for each industry and compare it with the two counterfactuals. We start by examining localization using the location-based counterfactual. We construct the confidence bands from 1,000 counterfactual K-density distributions of interfirm transaction relationship distances drawn from the pool of transaction relationships after controlling for the location of firms and their potential transaction partners. We do this exercise for each of the 150 three-digit manufacturing industries. For illustration, we present two figures showing the results for Ophthalmic Goods including Frames (JSIC316; Figure 4(a)) and Leather Footwear (JSIC214; Figure 4(b)).

The solid lines in these figures represent K-densities. Moreover, the bold dashed lines are the global confidence bands, while the thin dashed lines are the local confidence bands.

The figures show that the K-densities are overwhelmingly concentrated at the short end of distances in both industries.

Figures 4(a) and 4(b)

Further, we examine the share of manufacturing industries that can be classified either as localized or as dispersed in terms of interfirm transaction relationship distances relative to the location-based counterfactual. Figures 5(a) and 5(b) respectively depict the share of localized and dispersed industries in terms of their transaction relationships. Figure 5(a) shows that for short distances in the range of 0–40km transaction relationships of 90-95% of all manufacturing industries are localized. More specifically, most industries can be considered as localized in terms of their transaction relationships at relatively small scales, but the number of industries whose transaction relationships are localized falls rapidly at medium scales (70km). In contrast, Figure 5(b) shows that transaction relationships are dispersed only in a small number of industries at the smallest scales, while the number gradually increases at medium and large scales. The number of industries whose transaction relationships are dispersed within a range of 0–100km remains largely the same.

Figures 5(a) and 5(b)

4.3 Localization of Transaction Relationships Relative to Relationship-based Counterfactual

Next, we consider the localization of transaction relationships relative to the relationship-based counterfactual. We construct the confidence bands from 1,000 counterfactual K-density distributions of interfirm transaction relationship distances drawn from the pool of transaction relationships for the manufacturing industry as a whole.

We use the same two industries for illustrative purpose as in the preceding subsection, Ophthalmic Goods including Frames (JSIC316) and Leather Footwear (JSIC214). Figures 6(a) and (b) show the K-densities for the two industries, which are the same as in Figures

4(a) and (b), but the local and the global confidence bands are replaced with those generated by the location-specific counterfactual.

Figure 6(a) for the Ophthalmic Goods industry provides an example of interfirm transaction relationships being localized at the short end of distances. For every distance within the range of 0–50km, the K-density is above the upper global confidence band. Thus, the interfirm transaction relationships in this industry can be considered as localized in the range between 0–50km. On the other hand, Figure 6(b) for the Leather Footwear industry provides an example of transaction relationships being dispersed. For every distance within the range of 16–180km, the K-density is below the lower global confidence band and never above the upper global confidence band. Thus, interfirm transaction relationships in this industry are dispersed within the range.

Figures 6(a) and 6(b)

Next, we examine the share of manufacturing industries whose interfirm transaction relationships can be considered either as localized or as dispersed. Figures 7(a) and 7(b) respectively depict the share of localized and dispersed industries in terms of their transaction relationships. Figure 7(a) shows that for short distances in the range of 0-40km, the transaction relationships of almost 40% of all manufacturing industries are localized. The share of industries whose transaction relationships can be considered as localized falls rapidly for medium distances (40–60km), but gradually increases for long distances (60-180km). Note that most of the industries whose transaction relationships are localized at the large scales are different from industries whose relationships are localized at relatively small scales. On the other hand, Figure 7(b) shows that transaction relationships are dispersed in less than 20% of all manufacturing industries for the shortest distances, but the share gradually increases for medium distances, reaching a maximum of about 40% of all manufacturing industries at the distance of 70km. That share remains more or less unchanged for distances between 70 and 100km. Taken together the results shown in Figures 7(a) and (b) imply that, for the short distance range (0–40km), transaction relationships are more localized than the manufacturing industry average in some indus-

tries (about 40% of all manufacturing industries), while they are more dispersed in others (about 15-25% of them). Thus, although Figures 2(a) and 2(b) give the impression that there exists substantial inter-industry heterogeneity in terms of transaction relationship localization, we find that a considerable portion of industries (about 35-45%) are neither more localized nor more dispersed than the manufacturing sector as a whole.

Figures 7(a) and 7(b)

4.4 Industry Localization and Localization of Transaction Relationships

It has been widely argued that interfirm transactions are one of the most important reasons for industry localization (Rosenthal and Strange, 2001; Ellison, Glaeser, and Kerr, 2010). Therefore, in this subsection we empirically examine this view by focusing on the relationship between industry localization and the localization of interfirm relationships.

In order to illustrate how the two are related, we start by looking once again at the industries considered in the preceding subsections. In the previous subsection, we saw that transaction relationships in Ophthalmic Goods including Frames (JSIC316) were localized at small scales, while in Leather Footwear (JSIC214) they were dispersed. However, both of these industries were found to be localized in the study by Nakajima, Saito, and Uesugi (2010b), which applied the procedure proposed by Duranton and Overman (2005). Thus, taking the results obtained by Nakajima, Saito, and Uesugi (2010b) and those in this study together, it appears that while firms in Leather Footwear are located in close geographic proximity, their transaction relationships are dispersed. This suggests that, at least for some industries, industry localization and the localization of transaction relationships do not coincide.

Another way to look at the relationship between industry localization and the localization of transaction relationships is to compare the corresponding indices, that is, the index for the extent of industry localization (Γ'_A) and the index for the extent of localization of transaction relationships (Γ_A). Note that we employ the relationship-based

counterfactual for calculating Γ_A. Thus, Γ_A and Γ'_A respectively sum up the extent of departure from the overall tendency of transaction relationships and locations to be localized in manufacturing industry as a whole. Γ'_A sums up $\Gamma'_A(d)$ for all distances d, in which case the latter formula represents the difference between the K-density of all the pairwise distances of firms in a particular industry and the upper bound of the global confidence bands.[7]

In order to examine the relationship between the two indices, we calculate the correlation coefficient between the two and also plot a scattergraph. We find that the correlation coefficient is 0.25 and different from zero at a significance level of 1%, indicating that these two variables are positively correlated. This positive correlation coefficient is consistent with the results obtained by Rosenthal and Strange (2001) and Ellison, Glaeser and Kerr (2010) whose analysis is based on the premise that firms agglomerate in order to reduce the costs of transacting in goods and services. If firms indeed establish transaction relationships with other firms in the same agglomeration, their premise is likely to be correct. However, Figure 8 suggests that while there is a statistically significant positive correlation between Γ_A and Γ'_A, it is not overwhelmingly strong and there are a number of outliers, implying that their premise does not always hold.

Figure 8

Next, we examine whether and how the relationship between Γ_A and Γ'_A differs across industries. Tables 3 and 4 present the top twenty industries in terms of the highest values of Γ_A and Γ'_A, respectively. Several industries are included in both tables. These are Ophthalmic Goods including Frames (JSIC316), Physical and Chemical Instruments (JSIC314), Precious Metal Products and Jewels (JSIC321), Bookbinding and Printed Matter (JSIC163), and Industrial Plastic Products (JSIC193). Firms in these industries are not only geographically localized but also localized in their transaction relationships. While two of the industries (Opthalmic Goods and Physical and Chemical Instruments) fall into the precision instruments and machinery industry, there otherwise appears to be

[7] A detailed explanation of the definition of Γ'_A is provided in Duranton and Overman (2005) and Nakajima, Saito, and Uesugi (2010a).

no clear pattern in terms of which broader (two-digit) industries are more likely to show a strong positive association between Γ_A and Γ'_A. Further, Γ_A is zero and Γ'_A is positive in some industries, while the reverse is the case in others.

Tables 3 and 4

In Tables 3 and 4, there are seven industries falling into the former category, with a zero Γ_A and a positive Γ'_A: Cosmetics, Toothpaste and Toilet Preparations (JSIC177), Handbags and Small Leather Cases (JSIC217), Oil and Fat Products (JSIC175), Leather Footwear (JSIC214), Baggage (JSIC216), Electric Bulbs and Lighting Fixtures (JSIC273), and Miscellaneous Leather Products (JSIC219). Firms in these industries are geographically agglomerated but their transaction relationships are not localized. Four out of these seven belong to leather processing industries and two to chemical and allied products industries. We speculate that there are reasons for agglomeration other than the minimization of transaction costs with trading partners. For instance, in the leather processing industries, agglomeration may be due to the fact that, historically, those whose living was based on livestock processing, such as butchers or leather crafters, used to reside in segregated areas. If this is correct, the reduction of transaction costs with trading partners is not necessarily the primary reason for localization.

Next, there are six industries for which Γ'_A is zero and Γ_A is positive: Cement and Its Products (JSIC222), Paving Materials (JSIC184), Sliding Doors and Screens (JSIC143), Fabricated Constructional and Architectural Metal Products (JSIC254), Canned and Preserved Fruit and Vegetable Products (JSIC93), and Sawing, Planning Mills and Wood Products (JSIC131). All of these industries belong to different broader, two-digit industries. However, they do appear to share certain characteristics: their products are bulky (e.g., paving materials and wood products) or heavy (e.g., constructional metal products). Their suppliers are likely to be geographically dispersed since they are located in places where the natural resources or inputs they use are abundant, such as fruit and vegetables (for Canned Preserved Fruit and Vegetable Products) or limestone (for Cement and Its Products). Although firms in these industries tend to transact with counterparts in

close proximity to reduce the costs of dealing in heavy and bulky inputs and products, this does not provide sufficient incentive for agglomeration and firms instead locate close to suppliers that are geographically tied to a particular place due to natural resource considerations.[8]

In sum, there exists a positive association between Γ_A and Γ'_A, that is, the extent of localization of transaction relationships and the extent of industry agglomeration. However, it is important to emphasize that this is not the case for all industries. Some industries are geographically localized, while their transaction relationships are not. Conversely, in other industries, transaction relationships are localized, but firms in these industries are not agglomerated. Possible reasons for these exceptions include the historical background of a particular industry or location constraints that suppliers and their customers face.

5 Conclusion

This paper looked at the physical distances between transaction partner in Japanese manufacturing industry to examine the localization of transaction relationships. The findings can be summarized as follows. First, using the location-based counterfactual based on the location information of firms and their potential transaction partners, we find that about 90 to 95% of the 150 three-digit industries can be classified as localized in their transaction relationships at short distances. Put differently, this means that, given a number of potential transaction counterparts, firms in almost all industries tend to choose counterparts that are located in close proximity. This provides strong evidence that geographic proximity between transaction partners plays an important role in industrial agglomeration. Second, in about 40% of the 150 industries transaction relationships

[8]Note that the values of Γ_A and $\Gamma^{'}_A$ in the automobile and related industries are not very high and these industries do not appear in Table 3 or 4. Contrary to our expectation - alluded to in the introduction - that the automobile and related industries would be among the industries with the highest degree of localization and transaction relationship localization, there are actually many three-digit industries with higher values of $\Gamma^{'}_A$ and Γ_A. A possible explanation is that it takes a large amount of time and technical expertise for prospective auto parts suppliers to establish transaction relationships with automobile assemblers. Consequently, assemblers may tend to continue transacting with long-established suppliers even when these are located farther away.

are more localized than the relationship-based counterfactual representing the overall tendency for the localization of transaction relationships suggests. In other words, firms in these industries are more likely than the industry average to transact with firms in close proximity. And third, industrial agglomeration is positively associated with the localization of transaction relationships. However, some industries do not fit this pattern. For example, we find that leather processing industries tend to agglomerate in certain areas, but transaction relationships are not necessarily localized. On the other hand, in industries dealing with bulky or heavy products relying on local natural resource inputs (fruit and vegetables; limestone), firms tend to transact with partners in close proximity, without this providing sufficient incentive for agglomeration in these industries. In sum, transaction relationships are a key determining factor of agglomeration, but they are not the only one.

Using our dataset on firms' locations and their transaction relationships—something that is hard to come by not only in Japan but also elsewhere—several extensions of our analysis are possible. First, while our analysis, following the practice of previous studies, focused on manufacturing industries, our dataset makes it possible to extend the analysis to non-manufacturing industries. The nature of non-manufacturing sector activities means that transaction relationships are likely to differ considerably across industries. For example, the sector includes businesses as diverse as giant financial corporations and small mom and pop retail stores. Thus, the role that transaction relationships play in the diverse industries making up the non-manufacturing sector, the way they are related to the geographic localization of industries, and if so, how and why these patterns differ from the manufacturing sector are all potentially important research topics.

The second possible extension to the analysis here is to use the dataset in order to examine the relationship between transaction relationships and industry localization using a different approach. Here, we focused on bilateral transaction relationships and the distance between transaction partners in order to examine how this is related to industry localization. However, each bilateral relationship forms part of a bigger network of transactions. Hence, focusing on interfirm transaction networks rather than on individual

bilateral relationships and examining how such networks affect industrial agglomeration provides another interesting topic for further research. In fact, in another paper (Nakajima, Saito, and Uesugi, 2010b) we have already started examining interactions between transaction networks and the evolution of industry agglomeration, focusing on how network characteristics, such as the number of transaction relationships within a network and the distribution of such relationships, affect the extent of industrial agglomeration. Taken together, these studies can provide us with a deeper and more comprehensive understanding of industry agglomeration and interfirm relationships. Thus, the dataset employed here provides the scope for a range of studies that, taken together, should result in a better understanding of industry agglomeration and interfirm relationships.

References

Combes, P., T. Mayer, and J.F. Thisse (2008) *Economic Geography: The Integration of Regions and Nations*, Princeton University Press, Princeton.

Duranton, G. and H. Overman (2005), Testing for Localization Using Micro-geographic Data, *Review of Economic Studies* 72, 1077-1106.

Duranton, G. and H. Overman (2008), Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data, *Journal of Regional Science* 48, 213-243.

Duranton, G. and D. Puga (2004), Micro-foundations of Urban Agglomeration Economies, in V. Henderson and J. Thisse (eds.), *Handbook of Regional and Urban Economics* 4, 2063-2117, North-Holland, Amsterdam.

Ellison, G., E. Glaeser, and W. Kerr (2010), What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns, *American Economic Review*, 100(3), 1195-1213.

Murata, Y., R. Nakajima, R. Okamoto, and R. Tamura (2011), Localized Knowledge Spillovers and Patent Citations: A distance-based approach, *mimeo.*

Nakajima, K., Y. U. Saito, and I. Uesugi (2010a), Measuring Economic Localization: Evidence from Japanese Firm-level Data, *Journal of the Japanese and International Economies*, forthcoming.

Nakajima, K., Y. U. Saito, and I. Uesugi (2010b), The Role of Interfirm Transactions in Agglomeration and Coagglomeration: Evidence from Japanese Firm-level Data, *mimeo.*

Rosenthal, S. and W. Strange (2001), The Determinants of Agglomeration, *Journal of Urban Economics* 50, 191-229.

Silverman, B. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York.

Table 1: Top 10 industries with the smallest distances from transaction partners

JSIC	Industry name	Median	Mean	25 percentile	75 percentile	Standard deviation
105	Tobacco manufactures	4	93.81	3	199	132.98
169	Service industries related to printing trade	6	96.66	2	40	197.90
124	Japanese style apparel and "Tabi" socks	6	106.04	1	119	183.02
163	Bookbinding and printed matter	9	81.81	4	27	172.49
316	Ophthalmic goods including frames	10	108.78	3	161	173.99
162	Plate making for printing	10	120.29	3	149	209.38
161	Printing	12	120.98	4	140	212.43
216	Baggage	15	151.20	5	324	224.93
219	Miscellaneous leather products	16	170.57	5	402	228.78
321	Precious metal products and jewels	17	117.64	4	112	189.22

Note: All values are in kilometers.

Table 2: Top 10 industries with the largest distances from transaction partners

JSIC	Industry name	Median	Mean	25 percentile	75 percentile	Standard deviation
111	Silk reeling plants	335	333.91	63	541	266.97
117	Rope and netting	206	251.42	43	375	250.55
151	Pulp	169	271.92	17	495	283.07
174	Chemical fibers	163	191.21	12	321	184.00
121	Textile outer garments and shirts including bonded fabrics and lace non-Japanese style	139	213.28	9	389	233.81
202	Rubber and plastic footwear and its accouterments	137	222.58	7	399	253.32
112	Spinning mills	136	176.70	18	298	176.61
123	Underwear	135	208.50	14	390	216.04
122	Knitted garments and shirts	135	191.23	8	384	200.12
115	Knit fabrics mills	134	180.24	32	292	169.55

Note: All values are in kilometers.

Table 3: Top 20 manufacturing industries with the highest degree of industry localization

JSIC	Industry names	Γ' (Index of industry localization)	Γ (Index of transaction relationship localization)
316	Ophthalmic goods including frames	0.418	0.100
215	Leather gloves and mittens	0.332	0.005
314	Physical and chemical instruments	0.277	0.067
315	Optical instruments and lenses	0.272	0.035
321	Precious metal products and jewels	0.270	0.045
163	Bookbinding and printed matter	0.245	0.213
177	Cosmetics, toothpaste, and toilet preparations	0.235	0.000
112	Spinning mills	0.217	0.003
217	Handbags and small leather cases	0.213	0.000
113	Twisting and bulky yarns	0.213	0.017
118	Lace and other textile goods	0.210	0.028
175	Oil and fat products, soaps, synthetic detergents, surface-active agents, and paints	0.192	0.000
214	Leather footwear	0.190	0.000
216	Baggage	0.179	0.000
273	Electric bulbs and lighting fixtures	0.171	0.000
114	Woven fabric mills	0.170	0.015
224	Pottery and related products	0.165	0.000
193	Industrial plastic products	0.165	0.073
311	Measuring instruments, analytical instruments, and testing machines	0.165	0.018
219	Miscellaneous leather products	0.163	0.000

Note: Industries highlighted in blue have a Γ (Index of transaction relationship localization) of zero.

Table 4: Top 20 manufacturing industries with the highest degree of localization of transaction relationships

JSIC	Industry names	Γ' (Index of industry localization)	Γ (Index of transaction relationship localization)
163	Bookbinding and printed matter	0.245	0.213
161	Printing	0.063	0.111
116	Dyed and finished textiles	0.135	0.108
222	Cement and its products	0.000	0.105
255	Metal machine parts and tooling products	0.080	0.102
316	Ophthalmic goods including frames	0.418	0.100
256	Metal coating, engraving, and heat treating, excluding enameled ironware	0.057	0.099
184	Paving materials	0.000	0.078
259	Miscellaneous fabricated metal products	0.082	0.075
193	Industrial plastic products	0.165	0.073
314	Physical and chemical instruments	0.277	0.067
162	Plate making for printing	0.149	0.066
143	Sliding doors and screens	0.000	0.053
254	Fabricated constructional and architectural metal products including fabricated plate work and sheet metal work	0.000	0.049
93	Canned and preserved fruit and vegetable products	0.000	0.049
275	Electric measuring instruments	0.130	0.048
269	Miscellaneous machinery and machine parts	0.054	0.047
223	Structural clay products, excluding those of pottery	0.003	0.046
321	Precious metal products and jewels	0.270	0.045
131	Sawing, planning mills and wood products	0.000	0.045

Note: Industries highlighted in blue have a Γ (Index of industry localization) of zero.

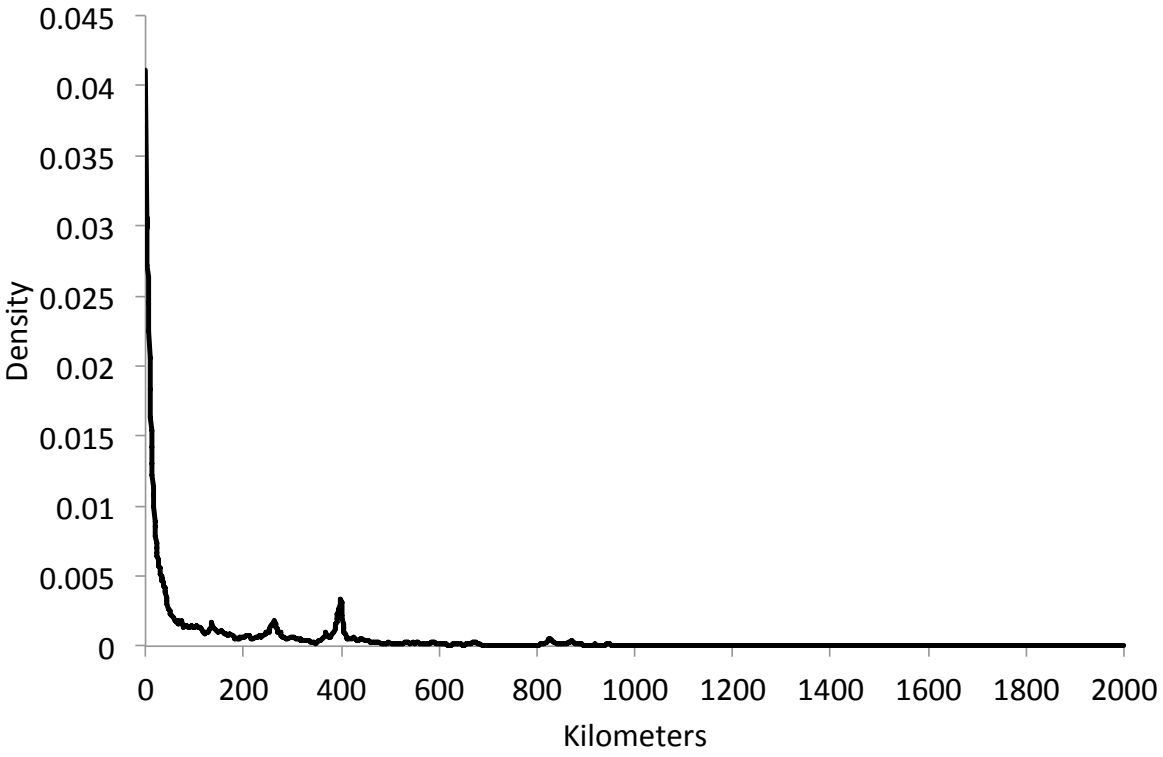
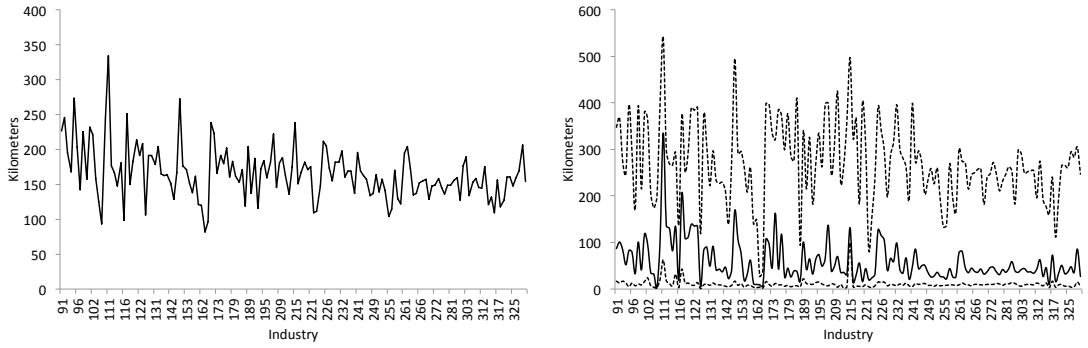
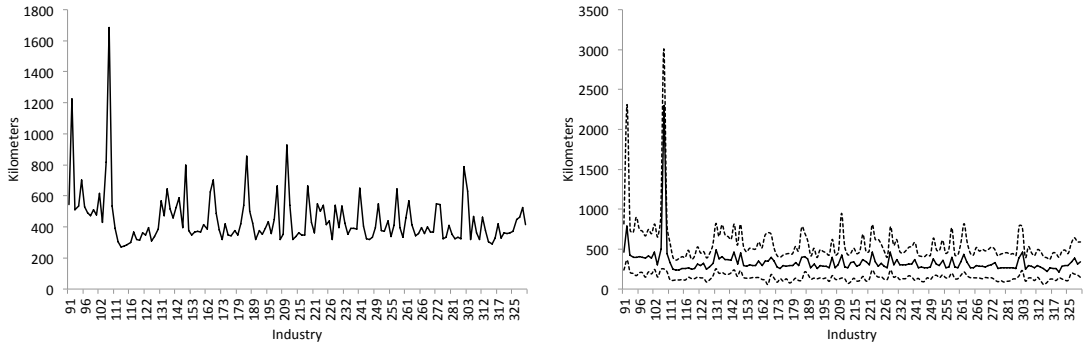


Figure 1: Probability distribution function of transaction relationship distances



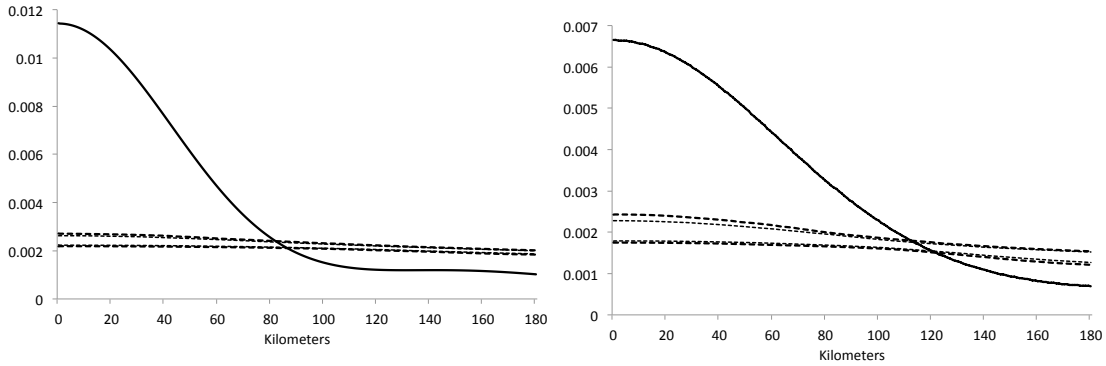
(a) Means (b) Quartiles

Figure 2: Summary statistics of transaction relationship distances



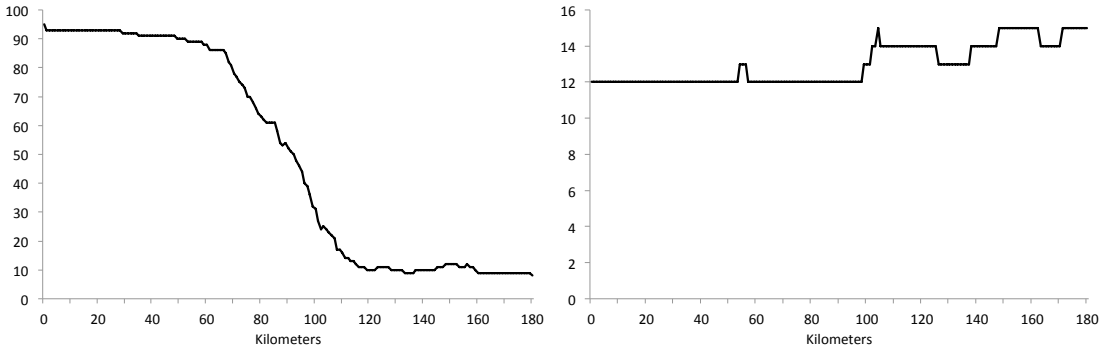
(a) Means (b) Quartiles

Figure 3: Summary statistics of transaction relationship distances between firms and their potential partners



(a) Ophthalmic Goods including Frames (JSIC316) (b) Leather Footwear (JSIC214)

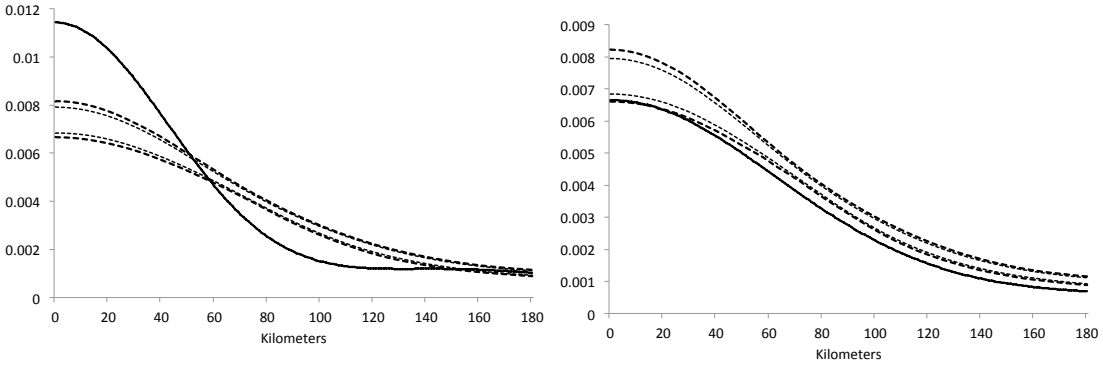
Figure 4: K-densities relative to the location-based counterfactual



(a) Share of localized industries

(b) Share of dispersed industries

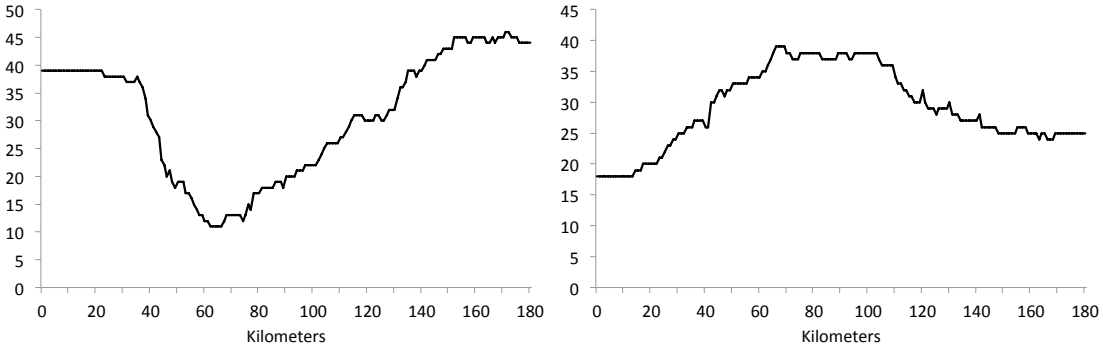
Figure 5: Share of localized and dispersed industries using the location-based counterfactual



(a) Ophthalmic Goods including Frames (JSIC316)

(b) Leather Footwear (JSIC214)

Figure 6: K-densities relative to the relationship-based counterfactual



(a) Share of localized industries

(b) Share of dispersed industries

Figure 7: Share of localized and dispersed industries using the relationship-based counterfactual

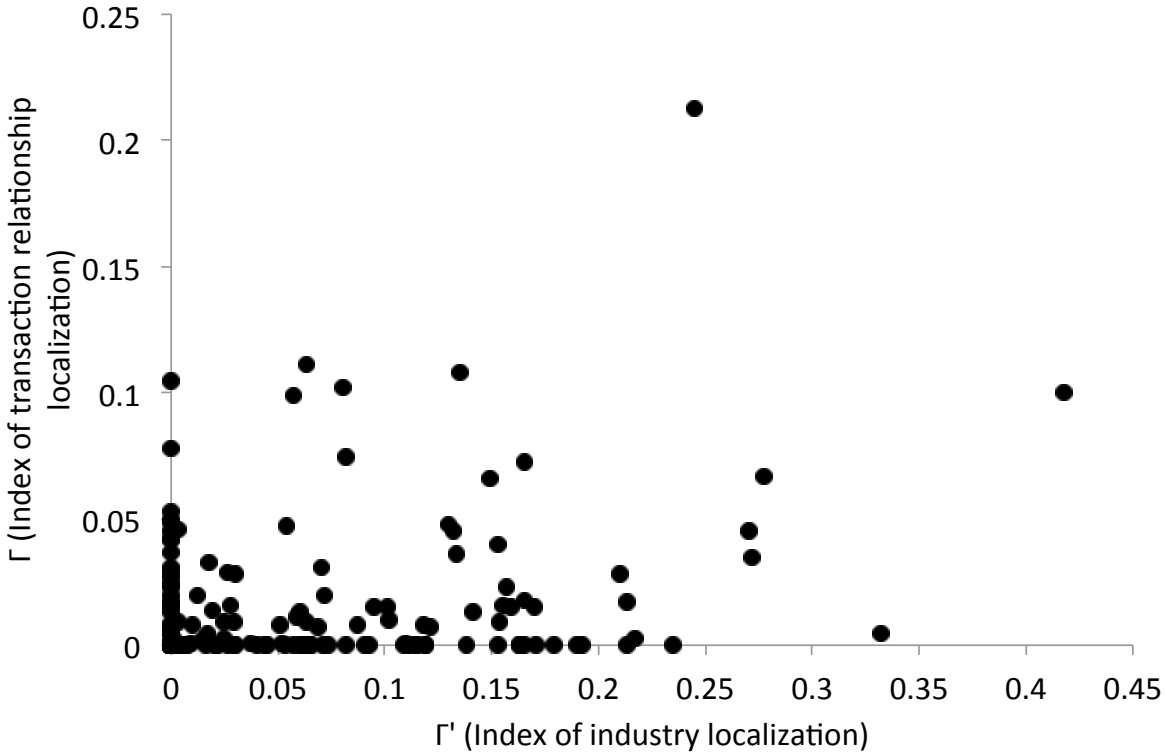


Figure 8: Relationship between Γ and Γ'