RIETI Discussion Paper Series 18-E-054

Inter-firm Transaction Networks and Location in a City

OTAZAWA Toshimori
Kobe University

OHIRA Yuki
Tottori University

Jos VAN OMMEREN
VU Amsterdam & Tinbergen Institute

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 18-E-054

Auggust 2018

Inter-firm Transaction Networks and Location in a City*

OTAZAWA Toshimori (Kobe University)

OHIRA Yuki (Tottori University)

Jos VAN OMMEREN (VU Amsterdam & Tinbergen Institute)

Abstract

This study contributes to the literature on the relationship between geographical and relation-based distances of economic agents. We aim to estimate the causal effect of a firm's position in the inter-firm transaction network on its spatial location within a city. Using micro data of inter-firm financial transactions for non-retail firms in the metropolitan areas of Japan, we demonstrate that the more central firms in transaction networks tend to have smaller inter-firm distances and therefore locate at more accessible places within the city. We also find that the results are robust to alternative specifications both of network centrality measures and spatial accessibility measures. It is also declared that the effect for single establishment firms are much stronger than that for multi-establishment firms. Furthermore, the result shows that this effect is noticeable for young firms in knowledge-intensive industries. The evidence suggests the potential importance of the inter-firm transaction pattern as a determinant of urban spatial configuration.

Keywords: Inter-firm transaction networks, Network centrality, Location proximity

JEL classification: L14, R12, D22

*This study is conducted as a part of the project "An Empirical Framework for Studying Spatial Patterns and Causal Relationships of Economic Agglomeration" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The author is grateful for helpful comments and suggestions by Tomoya Mori (Kyoto Univ.), Yukiko Saito (RIETI), and Discussion Paper seminar participants at RIETI. Otazawa gratefully acknowledges financial supports from the Grant-in-Aid for Challenging Exploratory Research (JSPS KAKENHI Grant Number JP 16K14320).

1 Introduction

How are spatial interactions and physical distance between firms interrelated? Our empirical approach to this issue is based on a theoretical perspective introduced by Helsley and Zenou (2014). Specifically, we aim to estimate the causal effect of relationship-based distance on geographical distance, using inter-firm transaction data for Japan.

Proximity reduces transport costs and facilitates interactions between economic agents. Marshall (1920) suggested that firms tend to locate near one another to reduce the costs of moving goods, people, and ideas. Since the pioneering work of Marshall, a number of studies have discussed the nature and sources of agglomeration economies from theoretical and empirical viewpoints.[1]

Spatial interactions underlie the formation of urban areas. Spatial externalities caused by market (financial transactions) and non-market interactions (e.g. knowledge spillover) between agents are likely a key determinant of urban spatial structure (see, e.g., Fujita (1988); Beckmann (1976)). Various models examine how these spatial externalities influence the location of firms and households, urban density patterns, and productivity.[2] These models assume that the extent of the externality attenuates to physical distance. Therefore, the level of the external effect that affects a particular firm (or worker) depends on the geographical location of agents and on the spatial arrangement of economic activities.

Helsley and Zenou (2014) is the first paper that explicitly incorporates the notion of social (relationship-based) distance into this issue and examines how interaction choices depends on the interplay of social and physical distance. Social distance depends on position of agents in social networks. Social networks are theoretical constructs useful in comprehending social interactions of a set of agents (such as individuals or firms), by depicting relationships among agents as graphs. Social networks have been studied by sociologists for more than a century and has grown into a central field of sociology (Wasserman and Faust (1994)). Until recently, most economists have been reluctant to consider social networks explicitly because of their analytical complexity. Last decade, however, studies of networks using game-theoretic modeling techniques and other economic perspectives have come to the fore. Ballester et al. (2006) establish a bridge between the economic and sociology social networks literature by relating behavior of

[1] For comprehensive reviews of this literature, see Duranton and Puga (2004) and Rosenthal and Strange (2004), respectively.

[2] Borukhov and Hochman (1977), O'Hara (1977), Tabuchi (1986), Lucus and Rossi-Hansberg (2002), Berliant et al. (2002), Helsley and Strange (2007) and Mossy and Picard (2011) are examples.

agents with their position in the social network and showing that the Nash equilibrium action of each agent is proportional to the Bonacich centrality measure, which is a commonly-used statistic of social network in sociology. More recently, Helsley and Zenou (2014) introduce geographical space into this model to examine how agents' interaction choices depend on the interplay of social and physical distance. They show that there is a tendency for those who are more central in the social network to locate closer to the geographical center of urban area. This theoretical perspective is based on the following two ideas; (i) saving transportation costs for interactions are an incentive for agents to locate closely each other, and (ii) agents in more central positions of social network can enjoy more benefit from location proximity. We aim to test the result by estimating the effect of the firms' position in the inter-firm transaction network on spatial location within a metropolitan area of Japan, using (non-retail) firm-level transaction data. An instrumental variables (IV) approach is applied to deal with endogeneity between firms' location choice and interaction choice.

Our paper fits in empirical literature on the relationship between inter-firm transactions and economic geography using inter-firm transactions micro data (see Nakajima (2015), for a review). For example, Bernard et al. (2015) examine the structure and geography of a domestic production network and its relationship with firm performance. Nakajima et al. (2013) demonstrate that the intensity of intra-industry transactions enforces industrial agglomeration. Itoh and Nakajima (2014) show that firm's centrality in a transaction network is an important determinant of FDI investment decision of Japanese firms.

The outline of the paper is as follows. In Section 2, we discuss the empirical strategy and introduce our measures of geographical distance and network centrality. In Section 3, we introduce the data and provide empirical results. Finally, Section 4 makes some concluding remarks.

2 Empirical strategy

In order to examine the effect of firms' position in the inter-firm transaction network on their geographical location within a city, we assume the following relationship between geographical location and the firms' position:

$$\ln(Access_i) = \alpha + \beta \ln(Centrality_i) + \sum_j \delta_j X_{i,j} + \epsilon_i, \tag{1}$$

where $Access_i$ is a geographical accessibility measure, $Centrality_i$ is a network centrality measure, $X_{i,j}$ refers to the other covariates and ϵ_i is an error term. As a geographical accessibility measure, we employ the entropy-based accessibility measure given as[3]:

$$eACC_i = \frac{1}{N} \sum_{j \neq i} n_j \cdot \exp(-\gamma \cdot dist_{ij}), \tag{3}$$

where $dist_{ij}$ is the Euclidean distance between firm i and j, n_j is the number of employees of firm j, $N = \sum_i n_i$ is the total number of employees in the city and $\gamma(> 0)$ is a distance decay parameter.

In the literature, there are a range of network centrality measures. We use PageRank centrality, PR_i, as a network centrality measure[4], which is defined as follows:

$$\begin{aligned} PR_i &= \sum_j \Big(\frac{1-\mu}{M} + \mu \frac{g_{ij}}{d_j} \Big) PR_j \\ &= \frac{1-\mu}{M} + \mu \sum_j \frac{g_{ij}}{d_j} PR_j, \quad \sum_j PR_j = 1. \end{aligned} \tag{4}$$

This measure is composed of two terms. The first term on the right hand side of (4), represents the effect of pure random connections to other firms. To include this effect is useful because it is plausible that firms have transactions with other firms with which they are not directly connected. μ is then the probability that firms have transaction with connected firms. M is total number of firms. Hence $(1 - \mu)/M$ is the probability of a random connection per firm. The second term represents the effect from network nodes with which firms are connected. It is based on the idea that if many important nodes j connect to node i, node i must have a higher centrality measure. To be specific, g_{ij} is an element of the adjacency matrix $\boldsymbol{G} = [g_{ij}]$, that keeps track of the direct connections in the network. Suppose $g_{ij} = 1$ if firm i is directly connected with firm j ; otherwise, $g_{ij} = 0$. We assume that if $g_{ij} = 1$ then $g_{ji} = 1$, so social links are reciprocal and set $g_{ii} = 0$. d_j is the number of firms with which node j connects, where $d_j = 1$ if node j has no connection. $1/d_j$ is an adjustment factor that evenly divides PageRank scores of j to nodes connected with j. So, if j has more relationships then i obtain less effects from being connected to j, all else being equal. This is thought of as an adjustment for the relative access or time that j can spend with i. PageRank centrality encompasses

[3]$eACC$ is one of basic specifications of the Hansen's measures, which is a widely-applied accessibility measure. The general definition of Hansen's accessibility measure is given as:

$$ACC_i = \frac{1}{N} \sum_{j \neq i} n_j \cdot f(dist_{ij}), \tag{2}$$

where $f(dist_{ij})$ is a damping function of distance.

[4]This measure has been devised originally by Google in order to rank web pages.

another well known network centrality measure, the degree centrality, which captures only the number of direct connections of a firm, when $\mu = 1$. We will assume that $\mu = 0.85$, so 15% of all connections are assumed to be random in the calculation of the PR.

In (1), we control for the firms' number of employees, capital stock, labor productivity, a limited company dummy and the access to transportation infrastructure, such as intercity railway stations and highway interchanges. In some specifications, we will also use industry fixed effects using a two-digit industrial classification. It controls for differences in land input, which is an important determinant of location choice, and for differences in production technology.

One may estimate (1) using OLS. A major concern with this is that the transaction relationship between firms may be the result of geographic location of firms, resulting in reverse causation. Some firms may trade with a firm because it is located close to them. Another, maybe more fundamental issue is that firms tend to stay at the same location for a long period, hence one aims to have information about (expected) centrality over a long period. In contrast, our measure of $Centrality_i$ is based on information about inter-firm transaction relationships for a certain limited period (one year), hence it is likely to have substantial measurement error, which usually results in downward-biased estimates. Finally, there are likely omitted variables that affect both the social network as well as location of firms (e.g., workforce characteristics such as educational level). Due to these issues, $Centrality_i$ may be endogenous as it is correlated to ϵ_i, which leads to a bias in the estimates. We deal with this issue using an instrumental variable (IV) approach. To estimate (1), we create an instrument using information of transactions with firms outside the city we focus on. The number of transactions outside the city strongly differs between firms. Hence, we argue that the number of transactions with firms outside the city affects the network centrality within the city, but does not directly affect the geographic location of firms within the city, when controlling the distance from terminal station and highway interchanges. In our estimations, we choose a metropolitan area[5] as a city. As it is regarded as economically integrated area in a sense which transactions within the area is much more than those outside the area, this seems a reasonable assumption. To capture this characteristic and to allow for non-linear effect of connections of firms outside the city, which is plausible because not all firms make any (substantial) transaction with firms outside the city, we employ the logarithm of the number of transactions with firms outside the city $\ln d_{out}$ as an instrumental variable.

[5]The definition of metropolitan area is described in 3.1.

We emphasize that this instrumental variable likely has much less measurement error than the $Centrality_i$, because they are based on information by a single firm.

In order to investigate the effects by industries, we also estimate the following specification:

$$\ln(Access_i) = \alpha + \sum_k \beta_k D_k^{icl} \ln(Centrality_i) + \sum_j \delta_j X_{i,j} + \epsilon_i, \tag{5}$$

where D_k^{icl} is a dummy variable that indicate 1-digit industrial classification k. $\ln(Centrality_i)$ is interacted with D_k^{icl} to allow distinct effects of network centrality by industries. For the estimation, we follow the approach of Balli and Sorensen (2013), which is to estimate a single first-stage to predict $\ln(\widehat{Centrality_i})$ using the instrument $\ln d_{out}$ and to use this predicted variable interacted with the dummies for industries in the second stage. This methodology is called Single-IV in this paper.

3 Data and results

3.1 Data and descriptives

We use firm-level micro data for the year 2014 extracted from a corporate profile database, that is compiled by a Japanese major credit research firm, Tokyo Shoko Research (TSR). The database contains standard profile of firms, i.e. name, location, industrial classification, date of foundation and registration, capital stock, number of employees, sales and profits. It also includes information about transaction relationships. Each firm reports its 12 main customers and 12 main suppliers. Hence, we have information on maximally 24 transaction relationships as reported by firms, because firms frequently report less than 24 relationships. A firm may also be linked as a customer or supplier to another firms which it does not report as a main customer or supplier, because other firm reports a relationship to this firm. Using this information, we can calculate a firms' transaction network, where each firm can be viewed as a node and a transaction between firms is represented by an edge between nodes. Because the database does not include all transactions among firms but only the main relationships, the network inevitably is a sub-network of the actual transaction network. Hence, the information is not perfect and will have substantial measurement error but will be shown to be sufficient to capture the structure of the inter-firm transaction network.

The main target area of this study is Hiroshima metropolitan area, that is defined based on census data in 2010. We employ '1.5% metropolitan areas' in the estimation. A metropolitan area is defined as one or more central cities and its associated peripheral

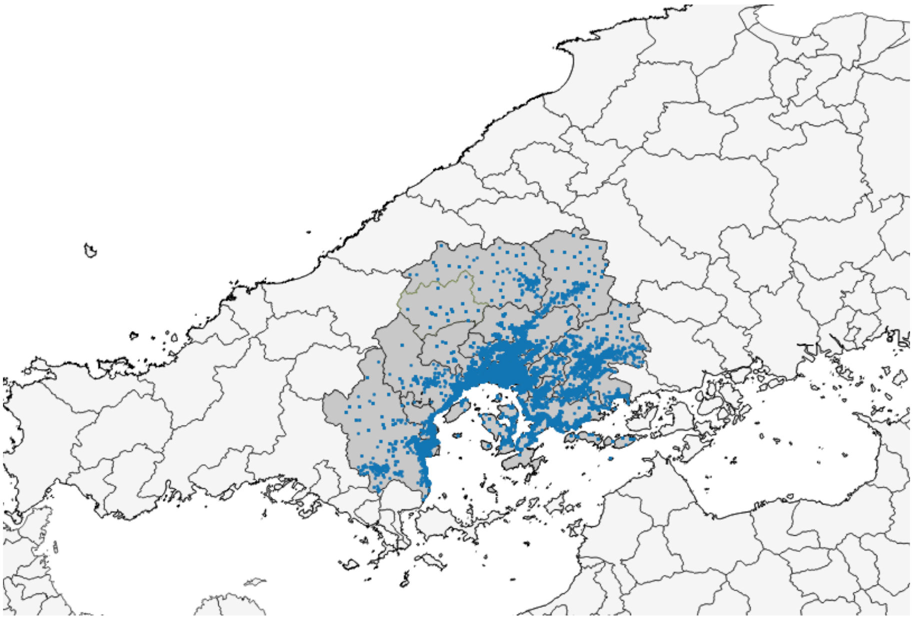


Figure 1: Location of firms in Hiroshima metropolitan area

municipalities. To qualify as a central city, a city must either be an ordinance-designated city or a non-designated city with a population of at least 500,000. To qualify as a peripheral municipality, the municipality must have at least 1.5% of its resident population aged 15 and above commuting to work or school to one of the central cities [6].

Hiroshima metropolitan area consists of Hiroshima city as the central city and 14 neighboring municipalities in Hiroshima prefecture and Yamaguchi prefecture as the peripheral area. A firm in Hiroshima metropolitan area has 6.03 transaction relationships within the city and 4.43 outside the city on average (see Table 2).

Figure 1 shows the location of firms within the city. About 88% of firms have more than one establishment in the city. For these firms, we have information about the location of the head quarter. Location choices of headquarters of multiple-establishment firms may differ from those of single-establishment firms.[7] For now, we show results in which these two kinds of firms are jointly analyzed. After that, we show the estimation

[6]There are many alternative definitions of metropolitan areas. The metropolitan area includes the built-up urban area and the economically connected territory to the outside. The economic relationship is generally defined by patterns of commuting to work into the urban area. Kanemoto and Tokunaga(2002) proposes urban employment area (UEA), which consists of a central city with at least 10,000 DID population and its surrounding municipalities whose 10% or more workers commute to the central city. UEA is typically smaller than and included in the 1.5% metropolitan area. We also conduct estimation applying UEA, but the results do not necessarily support our hypothesis. This implies that geography of local transaction networks may stretch more widely than the urban area measured by commuting patters. Identifying an appropriate geographical area for transaction-based industrial cluster is remained for future works.

[7]Behrens and Sharunova (2015) show that the characteristics of locations picked by multiunit plants differ in systematic ways from those picked by standalone plants.

Table 1: Basic information of Hiroshima metropolitan area

Population (a)	2,099,514 [people]		
Area (b)	5047.66 [km^2]		
Population density (a/b)	415.94 [people/km^2]		
	Full sample	Firms with transactions inside the area	Single establishment firms
Observations	29,439	18,136	2,884
Construction (D_4^{icl})	8,446	7,109	531
Manufacturing (D_5^{icl})	2,813	2,269	233
Information and communications (D_7^{icl})	512	235	38
Wholesale (D_9^{icl})	2,613	1,953	362
Scienteific research, professional and technical services (D_{12}^{icl})	1,578	697	93
Other industries	13,477	5,873	1,627

Table 2: Descriptives of key variables (Hiroshima metropolitan area)

VARIABLES	Mean	S.D.	Min.	Max
$eACC$	0.0217	0.0309	4.30e-06	0.152
$\ln(eACC)$	-4.996	1.756	-12.36	-1.882
$PR[\times 10^3]$	0.0543	0.100	0.0131	3.421
$\ln(PR)$	-10.18	0.731	-11.24	-5.678
no. of transactions with firms within the city	6.025	12.21	1	395
no. of transactions with firms outside the city	4.425	23.90	0	1,050
Capital [$\times 10^6$yen]	0.0480	2.490	1.00e-06	259.0
Employee	18.13	200.0	1	20,473
Labor productivity [$\times 10^6$yen per capita]	0.0540	2.807	0	306.2
Limited company dummy	0.536	0.499	0	1
dist. to terminal [$\times 10^2$m]	135.5	119.6	1.440	579.6
dist. to highway [$\times 10^2$m]	32.74	28.28	0.466	197.0

Note. No. of Observations : 12,263

result of single establishment firms and compare them. We calculate distance between all pairs of firms, $dist_{ij}$, based on the geocode data, by using Hubeny formula. We exclude certain industries in the analysis. We exclude retail and household service, accommodations, eating and drinking services as well as entertainments services, because the location of those services is mainly determined by the location of their consumers. In total, we employ 10,556 observations of firms in the benchmark estimation. We report the histograms of $eACC$ (for $\gamma = 0.1$) and PR in the Appendix. Descriptives of key variables are reported in Table 2.

3.2 Main results

Equation (1) and (5) have been estimated using ordinary least squares (OLS) as well as using instrumental variables (IV). Table 3 shows the main results. The OLS estimate of PageRank centrality elasticity is 0.0322 with a standard error of 0.0171. PageRank centrality elasticity estimated by IV is 0.508 (s.e. 0.0878), that is substantially higher than the OLS elasticity estimate.[8]

To interpret the size of the latter effect, it is useful to consider a one standard deviation increase in PageRank centrality, which is 0.10×10^{-3} (see Table 2). This increase implies a increase of 0.302 standard deviation in $\ln(eACC)$, equivalent to 0.153 standard deviation in eACC.[9] This result supports the aforementioned hypothesis, that is, the more central a firm is in the inter-firm transaction network, the more geographically accessible place it is located at.

Comparing columns (1) and (2), we find the coefficient of the IV regression is larger than the OLS coefficient. It should imply that the bias due to the measurement errors in the explanatory variable is larger than that due to the reverse causality.

In column (3), we show the estimates by industry. All estimates of interaction terms are positive and significant. The coefficient of construction sector, information and communication sector and scientific research, professional and technical services are greater than the aggregated estimates of all five industries, while that of manufacturing and wholesale is smaller.

[8]We use an F-test to determine the strength of the instrument. It turns out that the instrumental is highly significant, with a F-value of more than 100.

[9]This value is calculated in the following manner;

$$\beta \times \frac{\ln(\overline{PR} + \sigma_{PR}) - \ln(\overline{PR})}{\sigma_{\ln(eACC)}} = 0.508 \times \frac{\ln(0.1543 \times 10^{-3}) - \ln(0.0543 \times 10^{-3})}{1.756} = 0.302,$$

where $\overline{PR}$ and σ_{PR} denote the mean and the standard deviation of PR respectively and $\sigma_{\ln(eACC)}$ denotes the standard deviation of $\ln(eACC)$.

Table 3: Main result: Hiroshima metropolitan area

$\ln(eACC)$	(1) OLS	(2) IV	(3) Single-IV
$\ln(PR)$	0.0322*	0.508***	
	(0.0171)	(0.0878)	
$D_4^{icl} \cdot \widehat{\ln(PR)}$			0.706***
			(0.0907)
$D_5^{icl} \cdot \widehat{\ln(PR)}$			0.348***
			(0.0893)
$D_7^{icl} \cdot \widehat{\ln(PR)}$			0.684***
			(0.140)
$D_9^{icl} \cdot \widehat{\ln(PR)}$			0.495***
			(0.0862)
$D_{12}^{icl} \cdot \widehat{\ln(PR)}$			0.570***
			(0.129)
$\ln(employee)$	0.0659***	-0.0697**	-0.0761***
	(0.0131)	(0.0281)	(0.0268)
$\ln(capital)$	0.0464***	-0.0120	-0.0160
	(0.0132)	(0.0176)	(0.0170)
$\ln(lp)$	0.0378***	-0.0554***	-0.0674***
	(0.0126)	(0.0212)	(0.0202)
LTD. Dummy	0.120***	0.140***	0.131***
	(0.0248)	(0.0259)	(0.0251)
dist. to terminal	-0.00846***	-0.00842***	-0.00841***
	(0.000123)	(0.000126)	(0.000122)
dist. to highway	-0.0162***	-0.0162***	-0.0162***
	(0.000505)	(0.000511)	(0.000504)
Industry FEs	Yes	Yes	Yes
F-test (1st stage)	-	114.61	114.61
R-squared	0.650	0.626	0.653

No. of Observations : 10,556

Robust standard errors in parentheses.

***,**,*: significant at 1%, 5%, 10%, respectively.

We will now briefly discuss the results for the control variables for the IV estimates. In columns (2) and (3), number of employees and labor productivity decrease $eACC$, whereas limited company dummy increases $eACC$. Those results imply that larger and more labor productive firms tend to be located at less accessible places in the city, and a limited company tends to locate in more accessible places.

Table 4: Results of sensitivity analysis

$\ln(eACC)$	(1) $\gamma = 0.1$	(2) $\gamma = 0.1$	(3) $\gamma = 0.05$	(4) $\gamma = 0.1$	(5) $\gamma = 0.2$
$D_4^{icl} \cdot \ln(\widehat{Centrality})$	0.339***	0.596***	0.621***	0.706***	0.643***
	(0.0435)	(0.0764)	(0.0713)	(0.0907)	(0.100)
$D_5^{icl} \cdot \ln(\widehat{Centrality})$	0.168***	0.298***	0.336***	0.348***	0.301***
	(0.0428)	(0.0751)	(0.0705)	(0.0893)	(0.0983)
$D_7^{icl} \cdot \ln(\widehat{Centrality})$	0.330***	0.572***	0.594***	0.684***	0.675***
	(0.0663)	(0.116)	(0.0990)	(0.140)	(0.169)
$D_9^{icl} \cdot \ln(\widehat{Centrality})$	0.238***	0.419***	0.456***	0.495***	0.483***
	(0.0414)	(0.0726)	(0.0670)	(0.0862)	(0.0966)
$D_{12}^{icl} \cdot \ln(\widehat{Centrality})$	0.282***	0.481***	0.523***	0.570***	0.494***
	(0.0611)	(0.107)	(0.0978)	(0.129)	(0.142)
$\ln(employee)$	-0.0925***	-0.0769***	-0.0950***	-0.0761***	-0.0223
	(0.0290)	(0.0269)	(0.0209)	(0.0268)	(0.0298)
$\ln(capital)$	0.00197	-0.0231	-0.0377***	-0.0160	0.0106
	(0.0155)	(0.0177)	(0.0133)	(0.0170)	(0.0190)
$\ln(lp)$	-0.0568***	-0.0715***	-0.0632***	-0.0674***	-0.0508**
	(0.0190)	(0.0207)	(0.0156)	(0.0202)	(0.0227)
LTD. Dummy	0.0850***	0.125***	0.0880***	0.131***	0.157***
	(0.0252)	(0.0250)	(0.0194)	(0.0251)	(0.0281)
dist. to terminal	-0.00848***	-0.00843***	-0.00942***	-0.00841***	-0.00718***
	(0.000122)	(0.000122)	(0.000106)	(0.000122)	(0.000131)
dist. to highway	-0.0160***	-0.0163***	-0.0154***	-0.0162***	-0.0134***
	(0.000505)	(0.000504)	(0.000394)	(0.000504)	(0.000566)
Centrality	*Betweenness*	*Degree*	*PR*	*PR*	*PR*
Industry FEs	Yes	Yes	Yes	Yes	Yes
F-test (1st stage)	73.57	134.63	114.61	114.61	114.61
R-squared	0.653	0.653	0.766	0.653	0.553

No. of Observations : 10,556
Robust standard errors in parentheses.
***,**,*: significant at 1%, 5%, 10%, respectively.

3.3 Robustness

We will now examine the sensitivity of the results.[10] Table 4 shows the results using alternative measures for *Centrality* and other specifications of geographical accessibility measures *eACC*. Columns (1)-(2) report the results of betweenness centrality and degree centrality, respectively and (4) shows that of PageRank for comparison. It appears that all the estimates of interaction terms are positive and significant. Columns (3) and (5) report the results of the estimation using Hansen's accessibility measures for different value of γ. In all cases, the effect of PR is positive and significant and tendency of the results are unchanged. Hence, the reported effect of PR centrality on firms' location is

[10] Descriptives and correlations of variable using in the sensitivity analysis are reported in Table A.1 and A.2 , respectively.

Table 5: Single establishment firms

$\ln(eACC)$	(1) Single	(2) Single&Multi
$D_4^{icl} \cdot \widehat{\ln(PR)}$	1.536***	0.706***
	(0.427)	(0.0907)
$D_5^{icl} \cdot \widehat{\ln(PR)}$	1.043**	0.348***
	(0.433)	(0.0893)
$D_7^{icl} \cdot \widehat{\ln(PR)}$	1.423***	0.684***
	(0.520)	(0.140)
$D_9^{icl} \cdot \widehat{\ln(PR)}$	1.096***	0.495***
	(0.421)	(0.0862)
$D_{12}^{icl} \cdot \widehat{\ln(PR)}$	1.527***	0.570***
	(0.532)	(0.129)
$\ln(employee)$	-0.334***	-0.0761***
	(0.127)	(0.0268)
$\ln(capital)$	-0.129**	-0.0160
	(0.0607)	(0.0170)
$\ln(lp)$	-0.212**	-0.0674***
	(0.0985)	(0.0202)
LTD. Dummy	0.283***	0.131***
	(0.0829)	(0.0251)
dist. to terminal	-0.00949***	-0.00841***
	(0.000342)	(0.000122)
dist. to highway	-0.0159***	-0.0162***
	(0.00152)	(0.000504)
Industry FEs	Yes	Yes
F-test (1st stage)	20.32	114.61
R-squared	0.708	0.653
Observations	1,204	10,556

Robust standard errors in parentheses.
***,**,*: significant at 1%, 5%, 10%, respectively.

robust with respect to the specification of the accessibility measure.

3.4 Properties of firms

Location choice of a firm may depend on whether it is a single-establishment firm or a multi-establishment firm, or whether it is young or matured. Location of the headquarter of a milti-unit establishment is influenced by the spatial configurations of other units of the firm. Location of mature firms may be affected by the location of past transaction partners and not be optimal any longer. Therefore, accessibilities of single-unit and young firms are supposed to be more directly influenced by their transaction relationships than multi-unit firms and matured firms.

Column (1) in Table 5 shows the results of single establishment firms. The benchmark estimation in column (2) includes both of single- and multi-establishment firms. It

Table 6: Young firms

$\ln(eACC)$	(1) Young firms	(2) All
$D_4^{icl} \cdot \widehat{\ln(PR)}$	0.617	0.706***
	(0.403)	(0.0907)
$D_5^{icl} \cdot \widehat{\ln(PR)}$	0.320	0.348***
	(0.436)	(0.0893)
$D_7^{icl} \cdot \widehat{\ln(PR)}$	1.842**	0.684***
	(0.716)	(0.140)
$D_9^{icl} \cdot \widehat{\ln(PR)}$	0.358	0.495***
	(0.386)	(0.0862)
$D_{12}^{icl} \cdot \widehat{\ln(PR)}$	0.890*	0.570***
	(0.507)	(0.129)
$\ln(employee)$	-0.0430	-0.0761***
	(0.0906)	(0.0268)
$\ln(capital)$	-0.0147	-0.0160
	(0.0314)	(0.0170)
$\ln(lp)$	-0.0375	-0.0674***
	(0.0793)	(0.0202)
LTD. Dummy	0.117	0.131***
	(0.125)	(0.0251)
dist. to terminal	-0.00847***	-0.00841***
	(0.000436)	(0.000122)
dist. to highway	-0.0179***	-0.0162***
	(0.00205)	(0.000504)
Industry FEs	Yes	Yes
F-test (1st stage)	28.76	114.61
R-squared	0.606	0.653
Observations	1,022	10,556

Robust standard errors in parentheses.
***,**,*: significant at 1%, 5%, 10%, respectively.

is declared that the effect of centrality on accessibility for single establishment firms are much stronger than that for multi-establishment firms in all five industry sectors.

The estimation result of young firms is reported in column (1) in Table 6 while the benchmark case is shown in column (2). In this paper, we use the word "young" to refer to firms which established after 2006. All estimates of interaction terms are positive. In addition, the coefficients of the information and communications sector and those of the scientific research, professional and technical services sector are also larger than those of the benchmark case and significant. Both sectors are knowledge intensive and communications between suppliers and customers play important rolls in their business. Face-to-face contact is the most effective way of minute communication, but it necessarily requires transportation. Saving transportation costs for face-to-face communications may be a strong incentive for firms in those sectors to locate closely each other. Firms

Table 7: Comparison between other cities (1.5% metropolitan areas)

$\ln(eACC)$	(1) Hiroshima	(2) Sapporo	(3) Niigata
$D_4^{icl} \cdot \widehat{\ln(PR)}$	0.706***	0.190***	0.620***
	(0.0907)	(0.0567)	(0.114)
$D_5^{icl} \cdot \widehat{\ln(PR)}$	0.348***	-0.115*	0.302***
	(0.0893)	(0.0654)	(0.108)
$D_7^{icl} \cdot \widehat{\ln(PR)}$	0.684***	0.215**	0.961***
	(0.140)	(0.0903)	(0.151)
$D_9^{icl} \cdot \widehat{\ln(PR)}$	0.495***	0.187***	0.415***
	(0.0862)	(0.0544)	(0.106)
$D_{12}^{icl} \cdot \widehat{\ln(PR)}$	0.570***	0.378***	0.697***
	(0.129)	(0.0865)	(0.162)
dist. to terminal	-0.00841***	-0.00985***	-0.00464***
	(0.000122)	(0.000121)	(8.90e-05)
dist. to highway	-0.0162***	-0.0101***	-0.0118***
	(0.000504)	(0.000411)	(0.000481)
$\ln(employee)$	-0.0761***	0.0231	-0.187***
	(0.0268)	(0.0194)	(0.0360)
$\ln(capital)$	-0.0160	0.0656***	0.00882
	(0.0170)	(0.0123)	(0.0206)
$\ln(lp)$	-0.0674***	-0.0351**	-0.0606**
	(0.0202)	(0.0168)	(0.0293)
LTD. Dummy	0.131***	0.171***	0.210***
	(0.0251)	(0.0206)	(0.0299)
Industry FEs	Yes	Yes	Yes
F-test (1st stage)	114.61	213.91	98.85
R-squared	0.653	0.615	0.490
Observations	10,556	13,175	7,323

Robust standard errors in parentheses.
***,**,*: significant at 1%, 5%, 10%, respectively.

with higher network centrality enjoy more benefit from location proximity.

In the above cases, we have only about 10% of total observations in the benchmark case. If we estimate a separate first-stage regression for each industry, the weak instruments problem may occur due to the small sample size of each estimation. The single-IV approach resolves this problem by estimating a single (pooled) first-stage to predict $\widehat{\ln(PR)}$ using the instrument $\ln d_{out}$ and to use this predicted variable interacted with the dummies of the industry sectors in the second stage.

3.5 Comparison with other cities

We conduct estimation of other two cities, Sapporo and Niigata, applying 1.5% metropolitan area to these cities. In all of three metropolitan areas, more than 60% of firms have transactions inside the metropolitan areas and firms belonging to five sectors – construction, manufacturing, information and communications, wholesale, and scientific research, professional and technical services – account for about from 63% to 72% of those firms (See Table A.3 that describes numbers of observations by industrial sectors in Sapporo, Niigata metropolitan areas in addition to Hiroshima).

Columns (1), (2) and (3) in Table 7 shows the estimation results of Hiroshima, Sapporo, and Niigata metropolitan areas, respectively. The result of Niigata is similar to that of Hiroshima, and all coefficients of interaction sectors are positive and significant. On the other hand, the result of Sapporo is quite different from those of other two cities. More precisely, the estimate of the manufacture sector is negative and significant [11]. This result may be because manufacturing firms of Sapporo city are land intensive and tend to locate in the suburbs area, compared to other cities.

4 Concluding remarks

In this study, we aim to provide insight into the relationship between geographical distance and social (relationship-based) distance. We examine the effect of the positions in the inter-firm transaction network on spatial location of non-retail firms within a city using firm-level transaction data in a Japanese city. Using micro data of inter-firm financial transactions for non-retail firms in Hiroshima metropolitan area, we document that positions of firms in the transaction network affect the spatial locations in a metropolitan area. More precisely, one standard deviation increase in PageRank network centrality implies an increase of 0.153 standard deviation in accessibility measure of firms. The results are found to be robust to alternative specifications both of network centrality measures and spatial accessibility measures.

Applying Single-IV approach that resolves the weak instruments problem and provides the estimates by industrial sector regardless of the small number of observations in each industry, we find that geographical locations of single-unit establishments are more likely to be affected by their transaction relationships since their location decisions are more affected by location of (potential) transaction partners than multi-unit firms, and positions in transaction networks of young firms in knowledge intensive industries absolutely influence their geographical locations.

This paper provides the first evidence of relationship between transaction networks

[11] The result of detailed analysis in which 2-digit industrial classification, D_k^{icm}, is used as a dummy variable in eq.(5) shows that the coefficients of manufacture of food, manufacture of beverages, tobacco and fees, manufacture of textile products, pulp, manufacture of paper and paper products, printing and allied industries and manufacture of chemical and allied products are negative and significant.

and firms' location in a city. The evidence suggests the potential importance of the inter-firm transaction pattern as a determinant of urban spatial configuration.

Acknowledgements

The survey was conducted as a part of our research project at Japan's Research Institute of Economy, Trade and Industry (RIETI). The authors appreciate administrative supports by RIETI. This study was also supported by the Grant-in-Aid for Challenging Exploratory Research (JSPS KAKENHI Grant Number JP 16K14320).

References

[1] Ballester, C., Calvó-Armengol, A., Zenou, Y. (2006). Who's Who in Networks. Wanted: The Key Player, *Econometrica*, 74, 1403-1417.

[2] Beckmann, M.J. (1976). Spatial Equilibrium in the Dispersed City, In Papageorgiou, G.J. (Eds.), *Mathematical Land Use Theory*, MA: Lexington Books, 117-125, Lexington.

[3] Behrens, K., Sharunova, V. (2015). Inter- and Intra-firm Linkages: Evidence from Microgeographic Location Patterns, *Mimeo*.

[4] Berliant, M., Peng, S.K., Wang. P. (2002). Production Externalities and Urban Configuration, *Journal of Economic Theory*, Vol.104, 275-303.

[5] Bernard, A.B., Moxnes, A., Saito, Y.U. Production Networks, Geography and firm performance, *Journal of political economies*, forthcoming (*National Bureau of Economic Research Working Paper*, No.21082, 2015).

[6] Bonacich, P. (1987). Power and Centrality: A Family of Measures, *American Journal of Sociology*, Vol.92, No.5, 1170-1182.

[7] Borukhov, E., Hochman, O. (1977). Optimum and Market Equilibrium in a Model of a City without a Predetermined Center, *Environment and Planning*, A9, 849-856.

[8] Duranton, G., Puga, D. (2004). Micro-foundations of Urban Agglomeration Economies. In Henderson, J.V., Thisse,J.-F. (Eds.), *Handbook of Regional and Urban Economics*, Vol.4, North Holland, 2063-2117.

[9] Fujita, M. (1988). A monopolistic competition model of spatial agglomeration: a differentiated product approach,*Regional Science and Urban Economics*, Vol.18, 87-124.

[10] Helsley, R.W., Strange, W.C. (2007). Urban Interactions and Spatial Structure, *Journal of Economic Geography*, Vol.7, No.2, 119-138.

[11] Helsley, R. W., Zenou, Y. (2014). Social Networks and Interactions in Cities, *Journal of Economic Theory*, Vol. 150, 426-466.

[12] Itoh, R., Nakajima, K. (2014). Impact of Supply Chain Network Structure on FDI: Theory and Evidence, *RIETI Discussion Paper Series*, 14-E-27.

[13] Kanemoto, Y., Tokuoka, K. (2002). Metropolitan Area Definition in Japan, *Journal of Applied Regional Science*, No.7, 1-15 (in Japanese).

[14] Katz, L. (1953). A New Status Index Derived from Sociometric Analysis, *Psychometrika*, Vol.18, 39-43.

[15] Lucas, R.E., Rossi-Hansberg, E. (2002). On the Internal Structure of Cities, *Econometrica*, Vol.70, 1445-1476.

[16] Marshall, A. (1920). *Principles of Economics*, Maxmillan, London.

[17] Mossay, P., Picard, P.M. (2011). On Spatial Equilibria in a Social Interaction Model, *Journal of Economic Theory*, Vol.146, 2455-2477.

[18] Nakajima, K. (2015). Economic Geography and Interfirm Transaction Networks, In Watanabe, T. et.al. (Eds.), *The Economics of Interfirm Networks*, Advances in Japanese Business and Economics 4, 95-106, Springer, Japan.

[19] Nakajima, K., Saito, Y.U., Uesugi,I. (2013). Role of Inter-firm Transactions on Industrial Agglomeration: Evidence from Japanese Firm-level Data, *RIETI Discussion Paper Series*, 13-E-21.

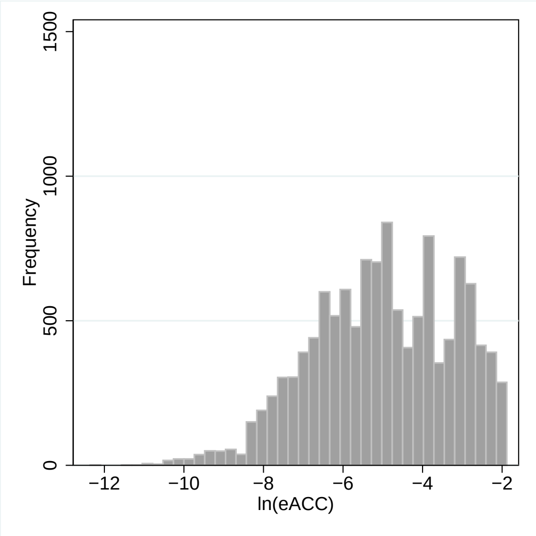
[20] O'hara, J. (1977). Location of Firms within a Square Central Business District, *Journal of Political Economy*, Vol.85, No.6, 1189-1207.

[21] Rosenthal, S., Strange, W. (2004). Evidence on the Nature and Sources of Agglomeration Economies, In Henderson,J.V., Thisse,J.-F. (Eds.), *Handbook of Regional and Urban Economics*, Vol.4, North Holland, 2019-2171.

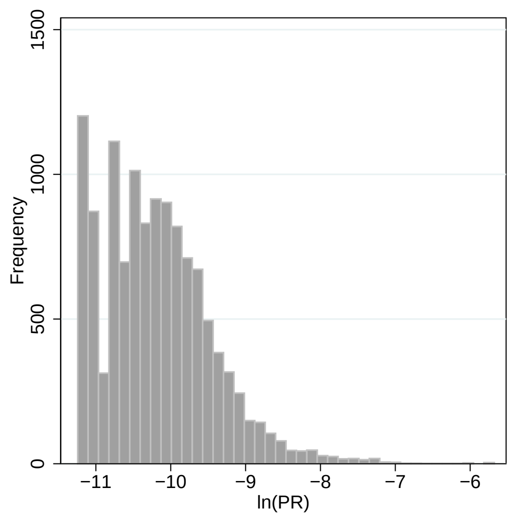
[22] Tabuchi, T. (1986). Urban Agglomeration Economies in a Linear City, *Regional Science and Urban Economics*, Vol.16, 421-436.

[23] Wasserman, S., Faust, K. (1994), *Social Network Analysis. Methods and Applications*, Cambridge University Press, Cambridge.

Appendix.1 Histograms of *Distance* and *Centrality*



(a) $\ln(eACC)$ $[\gamma = 0.1]$



(b) $\ln(PR)$ $[\mu = 0.85]$

Figure A.1 : Histograms of $\ln(eACC)$ and $\ln(PR)$ in Hiroshima metropolitan area

Appendix.2 Desctiptives and Correlations of Other Variables

Table A.1 : Descriptives of other variables

VARIABLES	Observations	Mean	S.D.	Min	Max
$e\text{–}ACC_1$ $[\gamma = 0.05]$	12,263	0.0576	0.0660	1.91e-05	0.246
$e\text{–}ACC_2$ $[\gamma = 0.1]$	12,263	0.0217	0.0309	4.30e-06	0.152
$e\text{–}ACC_3$ $[\gamma = 0.2]$	12,263	0.00701	0.0122	2.10e-06	0.0882
$\ln(e\text{–}ACC_1)$	12,263	-3.834	1.681	-10.87	-1.402
$\ln(e\text{–}ACC_2)$	12,263	-4.996	1.756	-12.36	-1.882
$\ln(e\text{–}ACC_3)$	12,263	-6.172	1.709	-13.07	-2.428
Betweenness	12,263	37,795	209,631	0	1.130e+07
Degree	12,263	6.025	12.21	1	395
PR	12,263	0.0543	0.100	0.0131	3.421
$\ln(Betweenness)$	9,488	9.050	2.086	-0.693	16.24
$\ln(Degree)$	12,263	1.265	0.940	0	5.979
$\ln(PR)$	12,263	-3.277	0.731	-4.335	1.230

Table A.2 : Correlations of variables

		e–ACC			ln(*e–ACC*)		
		$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.2$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.2$
e–ACC	$\gamma = 0.05$	1					
	$\gamma = 0.1$	0.9602	1				
	$\gamma = 0.2$	0.8460	0.9538	1			
ln(*e–ACC*)	$\gamma = 0.05$	0.8288	0.7191	0.5927	1		
	$\gamma = 0.1$	0.8603	0.7804	0.6667	0.9732	1	
	$\gamma = 0.2$	0.8394	0.7993	0.7183	0.9001	0.9693	1

No. of Observations: 12,262

	Betweenness	*Degree*	*PR*	ln(*Betweenness*)	ln(*Degree*)	ln(*PR*)
Betweenness	1					
Degree	0.8212	1				
PR	0.8104	0.9885	1			
ln(*Betweenness*)	0.3339	0.4585	0.4404	1		
ln(*Degree*)	0.4373	0.6744	0.6338	0.778	1	
ln(*PR*)	0.4749	0.7027	0.68	0.7937	0.9717	1

No. of Observations: 9,487

Appendix.3 Observations in Other Metropolitan Areas

Table A.3 : Observations by industrial sector in metropolitan areas

1.5%Metropolitan Area (MA)	Hiroshima	Sapporo	Niigata
Full sample in MA	29,439	35,513	20,462
Firms with transactions inside the MA	18,136	22,172	12,286
（by industrial sector）			
Construction (D_4^{icl})	7,109	7,453	4,270
Manufacturing (D_5^{icl})	2,269	1,959	2,399
Information and communications (D_7^{icl})	235	592	101
Wholesale (D_9^{icl})	1,953	2,830	1,629
Scientific research, professional and technical services (D_{12}^{icl})	697	1,308	339
Other industries	5,873	8,030	3,548