KEIO/KYOTO JOINT
GLOBAL CENTER OF EXCELLENCE PROGRAM
Raising Market Quality-Integrated Design of "Market Infrastructure"

KEIO/KYOTO GLOBAL COE DISCUSSION PAPER SERIES

DP2012-042

Agglomeration effects of inter-firm backward and forward linkages: evidence from Japanese manufacturing investment in China

Nobuaki Yamashita*
Toshiyuki Matsuura**
Kentaro Nakajima***

Abstract
This paper examines the agglomeration effects of multinational firms on the location decisions of first-time Japanese manufacturing investors in China for the period 1995–2007. This is accomplished by exploiting newly constructed measures of inter-firm backward and forward linkages formed in a home country. The conditional and mixed logit estimates reveal that agglomeration by first-tier suppliers and customers draws subsequent investment into a location. However, such agglomeration effects are not pervasive and do not extend to the second and third tiers. Instead, we find that agglomeration by third-tier suppliers generates a countervailing force, making a location relatively unattractive.

*Nobuaki Yamashita
La Trobe University, Melbourne
**Toshiyuki Matsuura
Keio University, Tokyo
***Kentaro Nakajima
Tohoku University, Sendai

KEIO/KYOTO JOINT GLOBAL COE PROGRAM
Raising Market Quality-Integrated Design of "Market Infrastructure"

Graduate School of Economics and Graduate School of Business and Commerce,
Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

Institute of Economic Research,
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Agglomeration effects of inter-firm backward and forward linkages: evidence from Japanese manufacturing investment in China[§]

Nobuaki Yamashita

La Trobe University, Melbourne

Toshiyuki Matsuura

Keio University, Tokyo

Kentaro Nakajima

Tohoku University, Sendai

Revised: 27/02/2013

Abstract:
This paper examines the agglomeration effects of multinational firms on the location decisions of first-time Japanese manufacturing investors in China for the period 1995–2007. This is accomplished by exploiting newly constructed measures of inter-firm backward and forward linkages formed in a home country. The conditional and mixed logit estimates reveal that agglomeration by first-tier suppliers and customers draws subsequent investment into a location. However, such agglomeration effects are not pervasive and do not extend to the second and third tiers. Instead, we find that agglomeration by third-tier suppliers generates a countervailing force, making a location relatively unattractive.

Key Words: Agglomeration, Backward and forward linkages, Location choice of multinational enterprises
JEL Classification: F23, L22, R3

[§] We acknowledge the financial assistance from the 21st century cultural and academic foundation in Tokyo, Japan. We are also grateful to Professors T. Tabuchi and K. Fukao for their support for this project. Thanks also go to Noel Gaston, Trevor Kollmann, and the workshop attendees at Kyushu University for their useful comments and suggestions.

1. Introduction

There is now ample evidence that multinational enterprises (MNEs) agglomerate in particular locations in a host country (see Head et al., 1995). For example, Debaere et al. (2010) reported that 60% of South Korean MNEs in the United States have located their manufacturing plants in the state of California, and 75% of them have established new affiliates in four provinces along the northeastern coast of China. Similar evidence is also found at a disaggregated geographical level in other host countries, such as France (Crozet et al., 2004), Portugal (Guimarães et al., 2000), and Italy (Roberto, 2004). It is commonly found that locations with many MNE plants belonging to the same industry or to vertically related industries are more likely to attract subsequent entries of MNE plants of the same national origin (Smith and Florida, 1994; Head et al., 1995, 1999; Head and Ries, 1996; Belderbos and Carree, 2002; Chang et al., 2013). This reflects the fact that the presence of MNE affiliates raises the probability of subsequent investment at the same location.

This paper investigates a new dimension of agglomeration effects of MNEs by considering *inter-firm* backward and forward linkages. Specifically, we examine the location decisions of Japanese manufacturing MNE start-ups across 22 Chinese provinces between 1995 and 2007.[1] We extend the idea that the presence of input-output (I-O) linkages of MNEs formed in a home country influences their co-location-cum-foreign direct investment (FDI) decisions in a host country. Moreover, the presence of inter-firm linked downstream or upstream affiliates draws further subsequent investment in particular regions due to cheaper access to existing suppliers and buyers.[2] This idea is not entirely new. Previous studies have

[1] In our dataset, the total number of Japanese MNE affiliates in China accounts for around 40% of total Japanese FDI worldwide.

[2] The importance of input-output (I-O) linkages in location choices of firms is highlighted by the New Economic Geography (NEG) models. Venables (1996) originally provided the theory of the interplay between vertically related industries and the forces of dispersion in the core-periphery economic structure. Subsequent work by Amiti (2005) considered vertically related industries under the conditions of various transportation costs and country asymmetry due to relative factor endowments. When industries are linked through an I-O structure, the downstream industry forms the market for upstream firms. To lower transportation costs, upstream firms are drawn to locations where there are relatively many downstream firms (backward linkages). Forward linkages suggest that a larger number of upstream firms located in one region can benefit downstream firms, which can obtain the intermediate inputs more cheaply by saving on transportation costs due to a large variety of differentiated inputs and more intense competition in upstream markets. These two vertical linkage effects motivate vertically related industries to cluster geographically.

tried to capture forward and backward industry linkages using I-O tables of a host country (Amiti and Javorcik, 2008) or a home country (Debaere et al., 2010), industrial groupings such as Japanese *keiretsu* (Head et al., 1995, 1999; Belderbos and Carree, 2002; Blonigen et al., 2005), or financial dependence (Mayer et al., 2010). However, these studies only explored the agglomeration effects of an immediate industrial relationship (what we term here as 'first-tier' linkages). We go much further. By capitalising on a unique feature of Tokyo Shoko Research (TSR) database, we identify the co-location of the first, second, and third tiers of multinational suppliers and customers, based on actual transaction-based records of *inter-firm* linkages. In this paper, we ask the following questions: how pervasive are the agglomeration effects by MNEs beyond the first-tier linkages? Do these effects vary at different tiers of inter-firm agglomeration? How do these results compare with those obtained from standard agglomeration measures?

It is important to consider the multiple layers of inter-firm linkages in the literature pertaining to MNE location decisions for the following reasons. First, such consideration can provide a much richer interpretation of the agglomeration effects of MNEs. As discussed by Mayer et al. (2010), the standard agglomeration variable for the stock of MNE affiliates operating in the same industry in a location can be quite broad since it represents various localisation economies. Our analysis considers both inter-firm backward linkages—the focus of previous studies—as well as the 'thickness' of the forward linkages. We find that the latter effect exerts comparatively stronger agglomeration effects. To our knowledge, only Debaere et al. (2010) considered both forward and backward linkages, although they used I-O tables. Additionally, our analysis shows that positive agglomeration externalities by inter-firm linkages are not pervasive and do not extend to the second and third tiers. Secondly, we can assess the relative strength of agglomeration effects in each layer—a first for a study of this kind. Our analysis finds new evidence of *negative* agglomeration externalities generated by the existence of third-tier suppliers in a location. This suggests that MNEs tend to avoid the same locations once the number of related input suppliers increases to 'too many'. However, no such effect was found for the agglomeration of related customers.

The next section introduces our method for the measurement of agglomeration by inter-firm backward and forward linkages. Section 3 describes the empirical implementation and the dataset used for the regression analysis. Section 4 discusses the results, and Section 5 concludes.

2. Measuring agglomeration

Agglomeration by inter-firm linkages

We employ the unique feature of the TSR database, which contains transaction information concerning inter-firm linkages among Japanese firms (Nakajima et al., 2012).[3] Section 3 provides detailed data descriptions. The basic idea is that we extract information of inter-firm linkages (such as which particular firm is linked with other firms through transactions concerning purchase and supply of outputs and inputs in vertical production chains forged in Japan) and then merge this information with FDI location choice data. In this way, we are able to track whether a supplier, for example, follows its customers by locating its foreign affiliate in the same location in a host country (known as the 'following-the-leader' type of FDI).[4] Presumably, location decisions of MNEs are influenced by the availability of intermediate input suppliers (backward linkage) and primal customers for their outputs (forward linkage) in a particular location/industry.

The original TSR file provides comprehensive coverage of inter-firm linkages with a maximum of 24 suppliers and customers for each individual Japanese firm. The TSR data traces, for example, a list of suppliers providing auto parts to the *Toyota Corporation* as well as a list of customers for *Toyota*'s outputs in the company's production chain. Inter-firm linkages extend within and across industries. This, of course, corresponds to traditional industry backward and forward linkages, but our measure captures this at the firm level. We use the original list to start tracing inter-firm linkages beyond the first-tier relations.

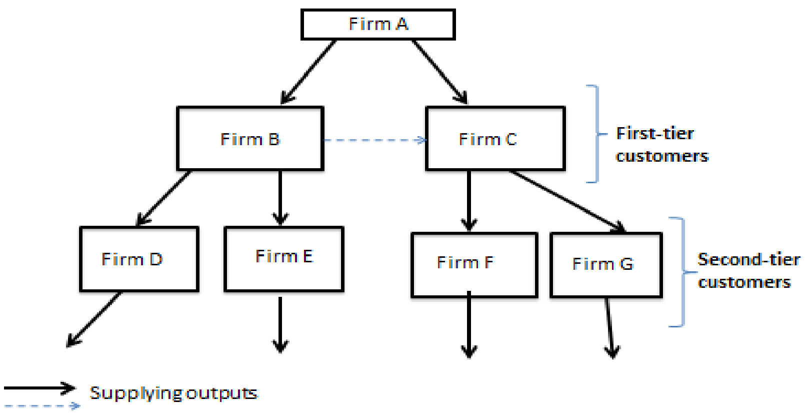
To explain the procedure, Figure 1 illustrates the hypothetical case of inter-firm forward linkages by Firm A supplying her outputs to Firms B and C, thus completing the first-tier transaction. We define Firms B and C as the first-tier customers from the viewpoint of Firm A. When Firms B and C supply their outputs to other firms, we denote the latter as the second-tier customers from the viewpoint of Firm A. We only define production chains in a unidirectional way. For example, as the dashed arrow indicates in Figure 1, if Firm B also

[3] TSR is Japan's credit reporting agency. It provides originally collected firm-level information pertaining to Japanese firms, for the purposes of corporate analysis. It also has a partnership with Dun and Bradstreet (D&B). However, to our knowledge, inter-firm linkages are not recorded in the worldwide version of the D&B database (Alfaro and Charlton, 2009).

[4] Naturally, for the purposes of our study, we only consider those suppliers and customers that have made FDI in China.

supplies to Firm C, we treat this as a new industrial linkage from the viewpoint of Firm B, with Firm C being the first-tier customer, and Firm A, the first-tier supplier. We repeat this exercise up to the third-tier linkages. Likewise, we define inter-firm backward linkages by identifying suppliers.

Figure 1: Inter-firm forward linkages (*FFL*) from the viewpoint of Firm A



Next, we sum up related suppliers and customers of different production in the same tier in a location to form agglomeration variables. It is crucial to note our algorithm for choosing firms to be scrutinised to alleviate the simultaneous location choices between the MNE's own affiliates and related affiliates. Consider Figure 1 once more. If, for example, Firm A established its affiliate in the year 2000, then we only count Firms B and C as related customers, as long as they had established their affiliates *before* 2000. In this way, we ensure a consistent unidirectional flow of production chains. When we move on to production chains from the viewpoint of Firm B, we drop Firm A, and then, we look at the establishment year of Firm C. More formally, the following inter-firm forward linkages (*FFL*) measure agglomeration of MNE *i* with location *p* invested in time *t*, by all related invested MNE affiliates (customers) *c* in time *s*.

$$FFL_{ipt}^{g} = \sum_{c} D_{cps}^{g} \quad \text{where} \quad g \in \{1,2,3\},\ s < t$$

where D represents a dummy variable equal to one for all related MNE affiliates c in location p created in year s, belonging to the g-th tier forward linkage from the viewpoint of MNE i.[5] Similarly, we construct a measure of inter-firm backward linkages (denoted as *FBL*) using counts of all related MNE affiliates (suppliers) in the g-th tier in the same location p at time s.

It is useful to point out the various strengths and limitations of our approach compared to previous studies. First, an inspection of a list of the first-tier relations reveals that *most of the relationships observed among* related suppliers and customers extracted from the TSR data are similar to those reported in the Japanese industrial grouping *keiretsu*'. Head et al. (1995, 1999), Belderbos and Carree (2002), and Blonigen et al. (2005) have already presented evidence for strong agglomeration effects generated within *keiretsu* groups. However, our agglomeration measures extend inter-firm relationships further, to the second and third tiers, which are not recorded for *keiretsu* group members. In addition, a *keiretsu* variable is usually confined to the relationship of vertically related supply firms (like the case of several major automobile manufacturers with associated auto input suppliers and electronics firms), whereas we also include the inter-firm relationships of purchasing firms.

As an illustration, Table 1 presents an example of actual inter-firm linkages of Toyota Motors, whose first affiliate in China was located in the Tienjin province in 1997. Table 1a shows the number of related suppliers[6] from the first to the third tier in all provinces. Table 1b breaks down information for first-tier suppliers, with an indication of whether the supplier belongs to Toyota's *keiretsu*. It seems that Tienjin province exerts strong agglomeration effects, because all three suppliers in the first tier belong to Toyota's *keiretsu* (Table 1b).

<Table 1 here>

Second, the inter-firm linkages in the TSR data do not show actual commodity flows with associated monetary values. This means all related suppliers and customers are treated unrealistically as being equally important. This differs from the data used by Holmes and Stevens (2012), namely, the values of commodity flows among U.S. establishments, sourced from the Commodity Flow Survey (CFS) of the U.S. Census Bureau. However, we are not

[5] The use of plant counts is standard in literature pertaining to FDI locations, while employment in plants is more frequently used in studies of regional and urban economics. The latter typically includes data of a much finer classification, including the numbers of manufacturing plants across various geographical locations.

[6] Note that we have concerned ourselves with suppliers only, because Toyota is primarily an assembler.

concerned about how important each individual supplier and customer is in relation to the location decision. Rather, we are more interested in the *relative importance* of clustered suppliers and customers in relation to location choices of MNEs. We can also reasonably assume that as we go further down (up) the layers of inter-firm transactions, after the first tier, lesser agglomeration effects would be generated. Hence, each layer should indicate the relative strength of I-O relationships. In addition, for our purpose, the unit of an investigation at the firm level is more appropriate, because all FDI decisions are made at the firm level rather than at the establishment level.

Third, inter-firm linkages are only considered for parent firms of a home country. This assumes that the same level of a technology and input requirements are carried over from parent firms of MNEs into their foreign affiliates. This is reasonable since an array of case study-based evidence, such as Moran (2011), suggests that foreign affiliates in a host country implement a similar technology to that employed in the home country.[7] In fact, there is evidence to suggest that Japanese MNEs (JMNEs) tend to replicate similar production chains both at home and in host countries (Belderbos and Sleuwaegen, 1996). More significantly, Barrios et al. (2011) also showed that I-O relations of a home country provide a closer approximation of sourcing behaviours of foreign affiliates than those of a host country.

Fourth, according to an analyst at TSR, inter-firm linkages only reflect the latest actual transaction information, which is constantly updated based on the more recent fieldwork surveys and follow-ups. This means that the time dimension of inter-firm linkages is defined at the time the data is accessed.[8] Accordingly, we assume inter-firm linkages are fixed during the period under study. However, it is well known that inter-firm relations in Japanese manufacturing remain relatively stable for a number of years, in fact, even as far as 10–15 years (Asamura, 1989). Also, note that studies using I-O tables at one point in time over log time intervals have made similar assumptions by arguing that the I-O relationship changes slowly over time (Mayer et al., 2010).

---

[7] Also see Yamashita (2010) for a detailed discussion of the link between Japanese parent firms and their foreign affiliates.

[8] We purchased the TSR data in 2010.

Agglomeration by industry linkages

Following previous studies (e.g. Head et al., 1995; Crozet et al., 2004; Debaere et al., 2010), we also introduce agglomeration measures within industry as well as industry linkages. Agglomeration measures within an industry in a location take two forms: the number of JMNE affiliates and the number of Chinese manufacturing plants. In line with the literature, we take into account of neighbouring agglomeration effects by computing the distance-weighted count of plants.[9] The agglomeration measure with the number of JMNEs (*NJ*) in a province *p* within industry *i* at time of investment *t* can be expressed by the following expression:

$$WI_{ipt} = NJ_{ipt} + \sum_{l \neq m} \frac{NJ_{imt}}{d_{mp}} \tag{1}$$

where *d* represents the bilateral distance between capital cities of provinces *p* and *m*. Eq. (1) suggests that *WI* will be higher if more JMNE affiliates exist in province *p* as well as the number of JMNE affiliates in nearby provinces, discounted by the relative distance to *p*. Alternatively, *NJ* can be replaced with the number of Chinese plants (*NC*) to reflect within-industry local agglomeration.

Additionally, similar to Debaere et al. (2010), we capture the industry-linkage dimension of agglomeration effects using the I-O tables of both the home country (Japan) and the host country (China). The Japanese I-O table is used to measure the agglomeration effects of industry backward linkages (*BL*) and forward linkages (*FL*), combined with the number of existing JMNE affiliates (*NJ*) in a location. At the same time, the Chinese I-O table is combined with the count of Chinese manufacturing plants in a region to capture the thickness of linkages to local industries (i.e. the availability of local suppliers and customers).[10] Presumably, the location decisions of MNEs are influenced by the availability of intermediate input suppliers (backward linkage) and primal customers for their outputs (forward linkage) in a particular location/industry. They are typically computed by constructing appropriate industry weights (a technical coefficient), as seen below.

[9] Note that the distance-weighted measures apply only to industry-level variables. They are not applicable to inter-firm agglomeration variables, because they have no variations within a location choice, thus making it impossible to estimate them.

[10] As mentioned in Debaere et al. (2010), we assume that a linkage with the local economy is reflected in the Chinese I-O table, though the count of plants in provinces may also include non-Chinese manufacturing plants.

$$W_{kj}^{B} = \frac{\text{inputs}_{k \leftarrow j}}{\text{total inputs}_k} \quad \text{and} \quad W_{kj}^{F} = \frac{\text{outputs}_{k \rightarrow j}}{\text{total outputs}_k}$$

W_{kj}^{B} is the share of inputs that industry k purchases from industry j in the total input purchases by industry k (superscript B indicates backward linkages). Conversely, W_{kj}^{F} is the share of outputs produced by industry k that are purchased by industry j from the total outputs produced by industry k (superscript F indicates forward linkages). Based on these two sets of industry weights, the following industry backward (BL) and forward agglomeration (FL) variables can be constructed in the case of JMNEs:

$$BL_{kpt} = \sum_{k,j} (W_{kj}^{B} \cdot NJ_{jpt}) \quad \text{and} \quad FL_{kpt} = \sum_{k,j} (W_{kj}^{F} \cdot NJ_{jpt})$$

These variables correspond to the weighted sums of the number of existing JMNE affiliates (NJ) in industry k in location p at the point of time t. The same formula can be applied in the case of Chinese plants (NC). We also take the neighbouring effects into account by using relative distance. A variable for distance-weighted backward linkages (WBL) with JMNE affiliates (NJ) is as follows:

$$WBL_{kpt} = BL_{kpt} + \sum_{k,j} W_{kj}^{B} \sum_{p \neq m} \frac{NJ_{imt}}{d_{mp}}$$

Distance-weighted forward linkages (WFL) can be constructed in a similar way.

In sum, we have two sets of within-industry agglomeration measures by Eq. (1) and four sets of inter-industry (backward and forward) linkage agglomeration measures for JMNE affiliates (NJ) and Chinese plants (NC).

3. Empirical implementation and data

We are primarily interested in identifying how regional variations of agglomeration influence an MNE’s choice to locate its first affiliates within China. We focus on the location decisions of first-time investors, because various locational attributes are perceived to be fixed at the time of investment. First, we implement the conditional logit model, which has been widely used since Head et al. (1995), for the problem of MNE location choice. The basic assumption of the model is that a firm (MNE) will choose to locate in the most profitable location, taking into account any positive externalities it can expect to receive (e.g. Japanese auto part suppliers are likely to locate near auto assembly plants) and other regional-specific attributes, such as the size of local demand and labour costs.[11] Basically, while we follow the model developed by Debaere et al. (2010), we use different notations to customise the problem.

Suppose that an underlying profit function for an MNE affiliate i choosing location p takes the following general form (for the time being, we omit an industry subscript).

(2) $$\pi_{\mathrm{ipt}} = \theta_p + A_{ipt}\alpha + Z_{\mathrm{pt}}\beta + \varepsilon_{\mathrm{ipt}}$$

where Z_{pt} is a vector of location-specific attributes varying by year of investment by MNE i, and θ_p denotes the time-invariant fixed effects of location attributes. If an MNE i selects a location p, then π_{ip} should be the highest among all alternative p choices. By assuming the type I value distribution in the error term in Eq. (2), the probability of MNE i choosing location p is expressed (without a time script t) as follows:

(3) $$\Pr(i \text{ locates in p}) = \frac{\exp^{\theta_p + \mathrm{A}_{ip}\alpha + Z_p\beta}}{\sum_m \exp^{\theta_m + \mathrm{A}_{im}\alpha + Z_m\beta}}$$

This can be estimated by the maximum likelihood estimation.

The most significant issue for the conditional logit estimation is the possible violation of the independence of irrelevant alternatives (IIA) assumption. The inclusion of the regional-specific effect, θ_p, in Eq. (2) provides a partial remedy, since it absorbs region-specific unobserved components in a profit equation. Further, we take the following approaches. First, we implement the mixed logit model estimation, which has been successfully applied in recent studies of the location choice problem by MNEs (Defever, 2006, 2012; Basile et al., 2008). This essentially allows for values of β with a subscript z to be random parameters, stemming from the heterogeneity of location choosers in Eq. (2). This can be expressed with the elements of the mean and deviation of β_z (note that in the case of the conditional logit model, the coefficients are fixed). Essentially, this allows the unobserved MNE affiliate

[11] More formally, when the production function of a firm is assumed to follow the Cobb-Douglas form, agglomeration externalities coupled with production inputs will affect the plant's output and profitability in a multiplicative way. In this case, the expected profitability in location can be expressed as a log-linear function of variables of the agglomeration effects and other locational attributes (Head et al., 1995).

characteristics to be correlated with the regional characteristics (Train, 2009).[12] Second, we report the results with the sub-samples by removing some regions or some groups of investors from the choice sets, to check the resilience of estimations. If the results remain unchanged even in the sub-samples, we can reasonably conclude that the IIA problem is not a matter of concern in our context (Head et al., 1995).[13]

At most, we have three sets of agglomeration variables: inter-firm backward and forward linkages (*FBL* and *FFL*), industry backward and forward linkages (*WBL* and *WFL*), and within-industry agglomeration measures (*WI*). Since each agglomeration measure captures different aspects of agglomeration elements, we try several specifications by including/excluding those agglomeration measures.[14] Other elements of regional attributes, Z_{pt}, in Eq. (2), include regional-specific manufacturing wages, market size, and policy incentive indicators. Market size at the province level are captured by the 'Harris' type of market potential (*MP*). This indicator is constructed in a fashion similar to Eq. (1), but using Gross Regional Products instead. As a proxy for a policy incentive indicator, we use the number of Special Economic Zones and Open Coastal Cities by province (*Economic Zones*). In the end, a fuller version of Eq. (2) can be written as follows:

$$\begin{aligned} \pi_{\mathrm{ipt}} = \theta_p + \sum_{g=\{1,2,3\}} \alpha_{1g} \ln FFL^g + \sum_{g=\{1,2,3\}} \alpha_{2g} \ln FBL^g + \sum_{\{NJ,NC\}} \alpha_3 WI \\ + \sum_{\{NJ,NC\}} \alpha_4 WFL + \sum_{\{NJ,NC\}} \alpha_5 WBL + Z_{\mathrm{pt}}\beta + \varepsilon_{\mathrm{ipt}} \end{aligned} \tag{4}$$

The most important coefficients are α_1 and α_2, which indicate the degree of agglomeration effects by inter-firm backward and forward linkages, while we control for other industry-level agglomeration effects.

[12] In the usual language of the mixed logit model, unobserved characteristics refer to heterogeneous taste parameters. As explained by Train (2009), while there are two interpretations of the mixed model, they may be formally considered as equivalent. However, it appears to us that the treatment of the mixed logit model in Basile et al. (2008) provokes an interpretation of 'error components', and that in Defever (2006, 2012), of the 'random coefficient'.

[13] In other words, this is a test of whether the ratio of probabilities of any two chosen alternatives is independent of all other alternatives.

[14] Another practical reason for doing this is that industry-level agglomeration variables show high correlations among themselves.

Data description

The main data is created by merging inter-firm linkages extracted from the TSR data with Japanese FDI data from *Overseas Japanese Companies Data* published by Toyo Keizai (TKZ).[15] We refer to the main data as the TSR-TKZ dataset. First, the sample of Japanese firms is confined to all Japanese manufacturing firms listed on the stock exchange in the TSR data (4,719 firms). Second, after extracting all inter-firm linkages by the procedure described in section 2, these firms are matched with the 2009 edition of the TKZ dataset, including the location of MNE plants across Chinese provinces with the 12 2-digit industry classifications, the year of establishment (since 1982), and the capital ownership ratio. At this stage, Japanese firms extracted from the TSR with no corresponding affiliates in China are excluded even if they have inter-firm linkages. As explained before, we only focus on the location choices of the first MNE plants between 1995 and 2007, although some MNEs have multiple plants established at several locations in China at different times. These filters reduce the number of location choosers down to 807 firms.

The cumulated number of JMNE affiliates at the industry level is sourced from the TKZ dataset and the *Global Reference Solution Database* (GRS) published by Dun and Bradstreet (D&B).[16] Counts of JMNE affiliates in Chinese provinces from the GRS are available from the 4-digit Standard Industrial Classification (SIC) System. We convert these into the 12 industry classifications in our main database. The key difference between the two data sources lies in the starting point of data recording for the entry of JMNE affiliates; 1992 for the GRS and 1983 for the TKZ. Therefore, the TKZ dataset covers a longer time period of history of Japanese investments in Chinese provinces for the same industry. However, we acknowledge that relying on a single database for measuring firm-level agglomeration, within an industry and inter-industry, may not be prudent. Hence, the GRS data is our preferred choice for industry-level agglomeration. However, both databases suffer from a common drawback: we do not have information for any affiliate exits. Once created, the agglomeration measures of JMNE affiliates kept adding up until the year 2007.

[15] The TKZ data is one of the most frequently used data sources for analyses concerning Japanese FDI (see Head et al., 1995; Belderbos and Carree, 2002).

[16] The GRS database is one of the commercial data products that form part of the *WorldBase* database of D&B. Alfaro and Charlton (2009) also employed the GRS database.

We use I-O tables from Japan and China for construction of industry-level agglomeration. The Japanese I-O table is sourced from the *Japan Industry Productivity Database* (Research Institute for Economy, Trade, and Industry).[17] We aggregate the matrix of 108 industries for 2007 into 12 2-digit industry classifications, in order to be consistent with the TKZ data.[18] The corresponding Chinese I-O table, provided by the Chinese Statistical Bureau, is also of the 2007 edition and contains a matrix of 12 industry flows. The annual average of the number of Chinese manufacturing plants and wages by province and industry are obtained from various years of the *China Manufacturing Statistical Yearbook* and the *China Labour Statistical Yearbook*, respectively. Gross regional domestic products (GRP) are sourced from various years of the *Chinese Statistical Yearbook*. Bilateral distances between provinces are calculated with the longitudes and latitudes of each province's capital city. We obtained the number of economic zones, including technology and industry development zones and export processing zones, from Table 2A.2 of Wang and Wei (2010).

4. Results

Table 3 reports the results by the conditional logit model and the mixed logit model in columns (1)-(3) and columns (4)-(6), respectively. The table shows some variations from the full model in Eq. (4). In column (1), we have inter-firm backward and forward agglomeration variables from the first to the third tier, together with within-industry agglomeration measures and regional fixed effects. While agglomeration variables, coupled with the within-industry agglomeration variable of JMNEs by first tier suppliers and customers, turn out to generate positive effects, this is not so for agglomeration variables by within-industry Chinese plants. This indicates that the location decisions of JMNEs are predominantly influenced by pre-existing locations of strongly linked JMNE affiliates as well as the general agglomeration of Japanese firms. We also note that, compared to backward linkages, inter-firm forward linkages show much stronger agglomeration effects in terms of magnitude. (This seems to be the case for all remaining regressions.) From the estimated coefficients, we can easily

[17] Available at: http://www.rieti.go.jp/en/database/JIP2012/index.html#04-1.

[18] It is standard practice to use a single year I-O table by assuming that a technical coefficient does not vary during the period under study.

interpret the average probability elasticity in the conditional logit model (Head et al., 1995).[19] The results in column (1) indicate that a 10% increase of the number of first-tier JMNE suppliers would lead to a 2% increase in the probability of a location being chosen by subsequent Japanese investors, whereas the agglomeration effect of first-tier customers is around 3.6%. Intuitively, this may mean that location choices by Japanese investors are relatively more influenced by the existence of relevant customers for their outputs in a location. This finding is indeed a generalisation of a study by Smith and Florida (1994), which showed that the location of large Japanese automobile assembly plants (customers) in U.S. states prominently drew subsequent investors of auto part suppliers to the same states. Our results further show that the desire to serve customers is relatively stronger than the desire to co-locate closer to suppliers.

More interestingly, the estimated coefficient of agglomeration by third-tier suppliers shows a negative sign with 1% statistical significance in Table 3. This suggests that agglomeration of inter-firm linked suppliers beyond the first tier can actually reduce the probability of a location to be chosen by subsequent Japanese investors. In the context of the New Economic Geography (NEG) model, this can be broadly interpreted as follows: as the density of economic activities of JMNEs in a location increases, the demands for factors of production (relevant workers or intermediate inputs) increase, resulting in rising production costs at this location and eventually, lower profits. Hence, rising factor prices by increased MNE locations become a countervailing factor, reducing agglomeration benefits discouraging co-locations. This ‘centrifugal’ force is an interesting aspect of inter-firm agglomerations, since the literature as a whole only shows positive agglomeration externalities.

Column (2) introduces a set of region-specific variables, instead of region fixed effects. The first notable change is that a coefficient for within-industry agglomeration of Chinese plants turns to a positive sign with a 5% statistical significance, whereas it was not statistically different from zero in column (1). Regional-specific variables (market potential and manufacturing wages) show marginal impact on the probability of location decisions. This could be a result of agglomeration variables capturing various localisation economics.

[19] This can be computed by the estimated coefficient multiplied by $(S - 1)/S$, where S stands for the number of alternative location choices (22 provinces in our case.)

In column (3), we add an industry linkage dimension of agglomeration to the model in column (1). Consistent with Debaere et al. (2010), we find that location choices of JMNEs are prominently driven by the availability of industry linkages along the same nationality (that is, with other JMNEs) and not with the local Chinese industries.[20] Again, we find that forward industry linkages with other JMNE affiliates are much more important than backward linkages.

We continue to use the mixed logit model in columns (4)-(6).[21] Overall, the results in the conditional logit model remain resilient. That is, we observe that while subsequent investments by JMNEs are prominently influenced by the agglomeration of first-tier suppliers and customers, these effects are not pervasive; at higher tier levels, supplier effects reduce the probability of a location to be selected by subsequent Japanese investors. While recent applications of the mixed logit model have proven it to be a powerful estimator capable of tackling the IIA problem (e.g. Defever 2006, 2012; Basile et al., 2008), we still believe that further checks by using sub-sample sets can be equally useful (Head et al., 1995; Debaere et al., 2010). In Table 4, we exclude three municipalities (Beijing, Tianjin, and Shanghai) from column (1), three Northeast provinces (Jilin, Liaoning, and Heilongjiang) from column (2), and Jiangsu (a province with the largest share of regional distributions of JMNE affiliates in Table 2) from column (3). Alternatively, we retain the full choice sets, but we exclude the automobile and electronics industries in columns (4) and (5). The location decisions of JMNEs in these two industries are eliminated, because they are usually characterised with an extensive coverage of vertical production processes. Again, the overall results generally remain unchanged, which is an encouraging sign from the viewpoint of the IIA assumption. Even in the limited choice sets, with the exception of column (3), we continue to observe strong agglomeration effects by first-tier suppliers and customers of JMNEs. Further supplier agglomeration in the third tier seems to reduce the probability of a location being chosen, as seen in Table 3. However, the first-tier supplier effect, in columns (3)-(5), becomes statistically insignificant, while the first-tier customer effect remains strong. The location decisions of many JMNE affiliates in the automotive and electronic industries (and their

[20] Although not indicated in Table 3, it is important to note that when we run experimental regressions with a variable of inter-industry linkages using only the Chinese I-O table, the linkage variables with Chinese industries are found to be positive and statistically different from zero. As soon as Japanese I-O information is included, these linkage variables become insignificant.

[21] Note that Table 3 does not report the estimated standard errors.

concentration in Jiangsu province) are overwhelmingly influenced by the existence of their first-tier suppliers in particular regions. This could be a result of earlier Japanese assembly factories from these industries having established themselves in China, followed by subsequent component suppliers. Further agglomeration becomes a countervailing factor for location choices of subsequent Japanese investors.

<Table 2>

<Table 3>

<Table 4>

5. Conclusion

There is an emerging consensus that the location decisions of MNEs are self-reinforcing, in the sense that subsequent investors are drawn to a location where many MNE affiliates (of the same nationality) from the same industry or from vertically related industries agglomerate. By extending this line of inquiry, we consider the agglomeration effects of *inter-firm* backward and forward linkages, using the case of first-time Japanese manufacturing investments across 22 Chinese provinces between 1995 and 2007.

Both the conditional logit and the mixed logit estimates revealed strong agglomeration effects exerted by first-tier suppliers and customers. The latter effect leads to a higher probability of a particular location being selected. We also found that the agglomeration effects are mainly driven by JMNEs in the electronics and automobile industries. This is consistent with the view that JMNEs like to replicate similar production chains that exist in Japan in a host country. At the same time, it was found that such agglomeration effects were short-lived and do not have any pervasive effects further down or up a vertical production chain. In fact, we found that agglomeration effects by third-tier suppliers actually lowered the probability that a location is selected by subsequent investors. We interpreted this as the increasing number of related suppliers making a location unattractive due to the rise in competition and increasing costs, such as the increased demand for inputs and labour. All in all, our measures of inter-firm agglomeration allow for a more detailed and nuanced interpretation of agglomeration effects attributable to MNEs.

References

Alfaro, L. and A. Charlton. (2009) 'Intra-industry foreign direct investment', *American Economic Review*, 99(5): 2096-2119.

Amiti, M. (2005) 'Location of vertically linked industries: Agglomeration versus comparative advantage', *European Economic Review*, 49(4): 809-832.

Amiti, M. and B. Javorcik (2008) 'Trade costs and location of foreign firms in China', *Journal of Development Economics*, 85: 129-149.

Asamura, B. (1989) 'Manufacturer-supplier relationships in Japan and the concept of relation-specific skill', *Journal of the Japanese and International Economies*, 3(1): 1-30.

Barrios, S., H. Görg and E. Strobl (2011) 'Spillovers through backward linkages from multinationals: Measurement matters!', *European Economic Review*, 55(6): 862-875.

Basile, R., D. Castellani and A. Zanfei (2008) 'Location choices of multinational firms in Europe: The role of EU cohesion policy', *Journal of International Economics*, 74: 328-340.

Belderbos, R. and L. Sleuwaegen (1996) 'Japanese Firms and the decision to invest abroad: Business groups and regional core networks', *The Review of Economics and Statistics*, 78(2): 214-220.

Belderbos, R. and M. Carree (2002) 'The location of Japanese investments in China: Agglomeration effects, *Keiretsu*, and firm heterogeneity', *Journal of the Japanese and International Economies*, 16: 194-211.

Blonigen, B. A., C. J. Ellis and D. Fausten (2005) 'Industrial groupings and foreign direct investment', *Journal of International Economics*, 65: 75-91.

Chang, K., K. Hayakawa and T. Matsuura (2013) 'Location choice of multinational enterprises in China: Comparison between Japan and Taiwan', *Papers in Regional Science* (forthcoming).

Crozet, M., T. Mayer and J. Mucchielli (2004) 'How do firms agglomerate? A study of FDI in France', *Regional Science and Urban Economics*, 34: 27-54.

Debaere, P., L. Joonhyung and M. Paik (2010) 'Agglomeration, backward and forward linkages: Evidence from South Korean investment in China', *Canadian Journal of Economics*, 43(2): 520-546.

Defever, F. (2006) 'Functional fragmentation and the location of multinational firms in the enlarged Europe', *Regional Science and Urban Economics*, 36, 658-677.

Defever, F. (2012) 'The spatial organization of multinational firms', *Canadian Journal of Economics*, 45(2): 672-697.

Guimarães, P., O. Figueiredo and D. Woodward (2000) 'Agglomeration and the location of foreign direct investment in Portugal', *Journal of Urban Economics*, 47(1): 115-135.

Head, K. and J. Ries (1996) 'Inter-city competition for foreign investment: Static and dynamic effects of China's incentive areas', *Journal of Urban Economics*, 40: 38-60.

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

Head, K., J. Ries and D. Swenson (1999) 'Attracting foreign manufacturing: Investment promotion and agglomeration', *Regional Science and Urban Economics*, 29: 197-218.

Holmes, T. J. and J. Stevens (2012) 'Exports, borders, distance, and plant size', *Journal of International Economics*, 88(1): 91-103.

Mayer, T., L. Mejean and B. Nefussi (2010) 'The location of domestic and foreign production affiliates by French multinational firms', *Journal of Urban Economics*, 68: 115-128.

Moran, T. H. (2011) *Foreign Direct Investment and Development: Launching a Second Generation of Policy Research*, Washington, DC: Petersons Institute for International Economics.

Nakajima, K., Y. Saito and I. Uesugi (2012) 'Measuring economic localisation: Evidence from Japanese firm-level data', *Journal of the Japanese and International Economies*, 26(2): 201-220.

Roberto, B. (2004) 'Acquisition versus greenfield investment: The location of foreign manufacturers in Italy', *Regional Science and Urban Economics*, 34(1): 3-25.

Smith, D. and R. Florida (1994) ‘Agglomeration and industry location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries’, *Journal of Urban Economics*, 36: 23-41.

Train, K. (2009) *Discrete Choice Methods with Simulation*, Cambridge: Cambridge University Press.

Venables, A. J. (1996) ‘Equilibrium locations of vertically linked industries’, *International Economic Review*, 37(2): 341-359.

Wang, Z. and S–J. Wei (2010) ‘What accounts for the rising sophistication of China’s exports’? in *China’s growing role in world trade*, eds. R. C. Feenstra and S–J. Wei (Chicago: University of Chicago Press).

Yamashita, N. (2010) *International Fragmentation of Production: The Impact of Outsourcing on the Japanese Economy*, Cheltenham: Edward Elgar.

Table 1a: Distribution of the number of related suppliers across China for Toyota Motors, whose first affiliate in China was located in the Tienjin province in 1997

State	First tier	Second tier	Third tier	Cumulated number	Province adjacent to Tienjin?	Located along the coast?
Shanghai	0	6	13	19		Yes
Jiangsu	3	3	13	19		Yes
Guangdong	0	3	14	17		Yes
Liaoning	0	2	10	12	Yes	Yes
Beijing	0	2	7	9	Yes	
Tienjin*	3	2	3	8		Yes
Shandong	1	1	4	6		Yes
Zhejiang	1	0	3	4		Yes
Fujian	1	0	3	4		Yes
Hebei	1	0	2	3	Yes	Yes
Hunan	0	1	2	3		
Shanxi	0	2	0	2		
Sichuan	0	1	0	1		
Guizhou	0	1	0	1		
Jilin	1	0	0	1		
Heilongjiang	0	0	1	1		
Jiangxi	0	1	0	1		
Hupei	0	0	0	0		
Hainan	0	0	1	1		Yes
Total	11	25	76	112		

Note: The entries are sorted in the descending order of the cumulated number of related suppliers. In each province, this number refers to all related suppliers that existed before 1997, the year Toyota Motors located its first affiliate in Tienjin province. Note that only related suppliers are identified, since the Toyota plant is usually the final assembler.
Source: TSR-TKZ data described in section 3.

Table 1b: List of first-tier suppliers to Toyota Motors and locations of their affiliates and year of establishment in China

State	Industry	Number of employees	Establishment year of affiliates	Toyota *keiretsu?*
Tienjin*	Transport equipment	1,497	1995	Yes
Tienjin*	Transport equipment	905	1994	Yes
Tienjin*	Transport equipment	1,480	1996	Yes
Hebei	Transport equipment	1,939	1996	Yes
Jilin	Rubber	-	1992	
Jiangsu	Transport equipment	292	1996	Yes
Jiangsu	Transport equipment	555	1996	Yes
Jiangsu	Rubber	-	1996	
Zhejiang	Transport equipment	685	1995	Yes
Fujian	Transport equipment	179	1995	Yes
Shandong	Transport equipment	431	1995	

Source: TSR-TKZ data described in section 3. The indicator for Toyota's *keiretsu* (see the final column in this table) is based on the TKZ data.

Table 2: Regional distribution of Japanese MNE (JMNE) affiliates and Chinese manufacturing plants in Chinese provinces for the period 1995-2007

Province	Coast?	JMNE affiliates[1]	Regional distribution	Chinese manufacturing plants[2]	Regional distribution
		Units	%	Units	%
East					
Beijing		23	2.3	6,219	3.0
Tienjin	Yes	51	5.1	5,569	2.7
Hebei	Yes	18	1.8	9,163	4.4
Shanghai	Yes	243	24.4	8,847	4.2
Jiangsu	Yes	244	24.5	23,324	11.2
Zhejiang	Yes	61	6.1	20,491	9.8
Fujian	Yes	17	1.7	7,042	3.4
Shandong	Yes	64	6.4	15,972	7.7
Guangdong	Yes	163	16.4	18,132	8.7
Guangxi		2	0.2	4,009	1.9
Liaoning	Yes	53	5.3	11,487	5.5
Hainan	Yes			494	0.2
Centre					
Shaanxi		1	0.1	3,873	1.9
Shanxi				3,617	1.7
Jilin		3	0.3	3,908	1.9
Anhui		6	0.6	7,004	3.4
Heilongjiang		2	0.2	4,786	2.3
Jiangxi		1	0.1	4,875	2.3
Henan		8	0.8	11,077	5.3
Hunan		6	0.6	7,682	3.7
Hupei		3	0.3	7,831	3.8
Neimenggu		3	0.3	2,527	1.2
West					
Sichuan		17	1.7	11,139	5.3
Guizhou		1	0.1	2,305	1.1
Yunnan		1	0.1	2,083	1.0
Qansu				2,476	1.2
Qinghai				474	0.2
Ninghsia		3	0.3	641	0.3
Xinjiang			0	1,515	0.7
TOTAL		994	100	208,560	100

Source*:* China Statistical Yearbook (various years), TKZ (2009).

Note*:* 1. Entries under this column refer to the cumulated number of first-time investments of JMNE affiliates between 1995 and 2008.

2. Entries under this column refer to the average number of Chinese manufacturing plants between 1995 and 2008.

Table 3: Location choices of first-time JMNE investors in China for the period 1995-2007

		Dependent variable: location choice					
		Conditional logit			Mixed logit		
		(1)	(2)	(3)	(4)	(5)	(6)
					Mean	Mean	Mean
Agglomeration by inter-firm forward or backward linkages	Variation across: Firm, province, and year						
First-tier *FFL* (Forward linkage by count of customers)		0.36***	0.38***	0.35***	0.41***	0.38***	0.35***
		[0.091]	[0.088]	[0.090]	[0.090]	[0.089]	[0.092]
Second-tier *FFL*		0.00	0.01	-0.03	0.04	0.01	-0.03
		[0.089]	[0.087]	[0.089]	[0.089]	[0.087]	[0.090]
Third-tier *FFL*		-0.04	-0.03	-0.08	-0.04	-0.03	-0.09
		[0.090]	[0.086]	[0.090]	[0.090]	[0.087]	[0.092]
First-tier *FBL* (Backward linkages by count of suppliers)		0.21*	0.24**	0.20*	0.26**	0.24**	0.20*
		[0.109]	[0.108]	[0.109]	[0.109]	[0.108]	[0.111]
Second-tier *FBL*		0.09	0.13	0.09	0.13	0.13	0.09
		[0.099]	[0.097]	[0.099]	[0.099]	[0.098]	[0.101]
Third-tier *FBL*		-0.26***	-0.22**	-0.25***	-0.23**	-0.22**	-0.25***
		[0.092]	[0.089]	[0.093]	[0.093]	[0.089]	[0.095]
Within-industry agglomeration	Industry, province, and year						
WI with *NC* (Count of Chinese plants)		0.12	0.21**	-0.09	0.15	0.23**	-0.10
		[0.114]	[0.103]	[0.319]	[0.120]	[0.107]	[0.325]
WI with *NJ* (Count of JMNE affiliates)		1.49***	1.09***	0.21	1.06***	1.11***	0.21
		[0.114]	[0.075]	[0.142]	[0.117]	[0.084]	[0.159]

		(1)	(2)	(3)	(4)	(5)	(6)
Agglomeration by industry linkages	Industry, province, and year						
WFL with *NC* (Forward linkages by Chinese plants)				-0.44			-0.34
				[0.596]			[0.622]
WBL with *NC* (Backward linkages by Chinese plants)				0.72			0.66
				[0.525]			[0.542]
WFL with *NJ* (Forward linkages by JMNE affiliates)				1.17***			1.14***
				[0.322]			[0.336]
WBL with *NJ* (Backward linkages by JMNE affiliates)				0.50			0.58*
				[0.316]			[0.334]
Region-specific variables	Province and year						
Market potential (*MP*)			0.29*			0.29*	
			[0.170]			[0.170]	
Manufacturing wages			0.29*			0.29*	
			[0.168]			[0.169]	
Economic zones			-0.10			-0.10	
			[0.131]			[0.131]	
Province dummy		Yes	No	Yes	Yes	No	Yes
Observations		17,710	17,710	17,710	17,710	17,710	17,710
Log-likelihood		-1601	-1657	-1612	-1646	-1657	-1611

Note:
1. We take log values for all variables except Economic Zones. Standard errors appear in parentheses.
2. Statistical significance (two-tailed test): ***, **, and * imply statistical significance at the 1%, 5%, and 10% level, respectively.
3. The 22 Chinese provinces make up the choice sets.
4. Standard errors of the random coefficients are not shown, since most of them are not statistically significant.
5. All variables also vary according to year of investments.
6. Industry-level variables are distance-weighted. See section 2 for variable construction.

Table 4: Sub-sample location choice sets by the conditional logit model

	Dependent variable: location choice				
	(1)	(2)	(3)	(4)	(5)
	Excluding:				
	Municipalities	Northeast	Jiangsu	Automobile	Electronics
Agglomeration by inter-firm forward or backward linkages					
First-tier *FFL* (Forward linkage by count of customers)	0.28**	0.36***	0.43***	0.30***	0.35***
	[0.118]	[0.094]	[0.109]	[0.102]	[0.101]
Second-tier *FFL*	0.04	-0.03	-0.05	-0.02	-0.06
	[0.115]	[0.092]	[0.102]	[0.100]	[0.099]
Third-tier *FFL*	-0.03	-0.10	-0.13	-0.03	-0.14
	[0.111]	[0.093]	[0.104]	[0.101]	[0.100]
First-tier *FBL* (Backward linkages by count of suppliers)	0.30**	0.25**	0.17	0.08	0.16
	[0.143]	[0.114]	[0.136]	[0.124]	[0.123]
Second-tier *FBL*	0.14	0.10	0.12	0.07	0.12
	[0.128]	[0.104]	[0.116]	[0.111]	[0.112]
Third-tier *FBL*	-0.34***	-0.27***	-0.16	-0.23**	-0.21**
	[0.113]	[0.097]	[0.108]	[0.104]	[0.104]
Within-industry agglomeration					
WI with *NC* (Count of Chinese plants)	-0.21	0.09	-0.05	0.09	-0.19
	[0.387]	[0.344]	[0.334]	[0.335]	[0.334]
WI with *NJ* (Count of JMNE affiliates)	0.10	0.29*	0.25*	0.04	0.32**
	[0.162]	[0.151]	[0.145]	[0.162]	[0.151]
Agglomeration by industry linkages					
WFL with *NC* (Forward linkages by Chinese plants)	-0.49	-0.85	-0.57	0.46	-0.19
	[0.719]	[0.640]	[0.626]	[0.702]	[0.616]
WBL with *NC* (Backward linkages by Chinese plants)	1.00	0.77	0.75	-0.29	0.62

	[0.622]	[0.565]	[0.543]	[0.687]	[0.547]
WFL with *NJ* (Forward linkages by JMNE affiliates)	1.02**	1.22***	1.12***	0.84**	1.11***
	[0.400]	[0.346]	[0.324]	[0.385]	[0.329]
WBL with *NJ* (Backward linkages by JMNE affiliates)	0.81**	0.39	0.45	0.83**	0.52
	[0.403]	[0.330]	[0.320]	[0.343]	[0.322]
Province dummy	Yes	Yes	Yes	Yes	Yes
Observations	10,526	14,269	12,957	15,092	14,168
Log-likelihood	-967.7	-1404	-1197	-1363	-1307

Note:
1. All variables are in log form.
2. Standard errors appear in parentheses.
3. Statistical significance (two-tailed test): ***, **, and * imply statistical significance at the 1%, 5%, and 10% level, respectively.