Self-Organization in the Spatial Economy: Size, Location and Specialization of Cities

Tomoya Mori* and Koji Nishikimi†

November 8, 2001

Abstract

The spatial distribution of industries and population is quite lumpy, and the lumpiness varies across industries as well as across different types of population depending on their education level. We show using Japanese data that by a certain choice of aggregation level for industries as well as region, and with an appropriate method for identifying industry location, the location of industries and that of population can be related by surprisingly simple and persistent patterns. In particular, our choice of regional and industrial categories turned out to be less aggregate than those often used in the previous literature. Despite the seemingly complex nature of geographical and industrial systems using less aggregate units, the resulting location patterns are much simpler than in a more aggregated system.

The latest version of this paper is available from
http://www.kier.kyoto-u.ac.jp/~mori/papers/mn01.pdf

Keywords: Industry location; Aggregation problem; Number, size, location and specialization of metropolitan areas; Critical population size for industries; Hierarchical principle; Agglomeration; Education level; Self-organization; Complex systems

Acknowledgment: We thank Duncan Black, Gilles Duranton, Masa Fujita, Robert Helseley, Hideo Konishi, Maria P. Makabenta, Se-il Mun, Komei Sasaki, Akihisa Shibata, Chiharu Tamamura, Tatsufumi Yamagata, Atsushi Yoshida, and seminar participants at Kyoto University, Tohoku University and Tokyo University for their constructive comments and encouragement, and Yoshi Kanemoto for providing us the processed data and algorithm for the establishment of the metropolitan areas in Japan. The research assistance by Yuko Takeda for the construction of intercity distance data, and Noriko Tabata for the construction of county data are greatly acknowledged. This research has been supported by The Murata Science Foundation, The Grant in Aid for Research (Nos. 09CE2002, 09730009, 13851002) of Ministry of Education, Science and Culture in Japan.

*Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501 Japan. Email: mori@kier.kyoto-u.ac.jp. Phone: +81-75-753-7121. Fax: +81-75-753-7198.

†Institute of Developing Economies, 3-2-2 Wakaba, Mihama-ku, Chiba, 261-8454 Japan. Email: nisikimi@ide.go.jp. Phone: +81-43-299-9672. Fax: +81-43-299-9763.

1 Introduction

The spatial distribution of industry and population is quite lumpy, and the lumpiness varies across industries as well as different types of population in terms of education level. From the commuting pattern data of workers, we can see that economic activities are concentrated in a small part of the national location space. In 1995, 80% of the workforce in Japan are found in metropolitan areas[1] which occupy only 30% of the geographic space in the nation. The distribution of industries and population in the metropolitan areas is far from uniform. Among the 149 three-digit manufacturing industries, 45 [resp., 22] have positive employment in less than 50% [resp., 25%] of the metropolitan areas, while 42 are ubiquitous, having positive employment in more than 90% of the metropolitan areas. Of the total population in the metropolitan areas, more than 60% is concentrated in the largest ten, and more than 30% reside in Tokyo. The education levels of workers in metropolitan areas vary widely as well. While the workers with high school diplomas are almost equally distributed (in terms of share in total workforce) in each metropolitan area, those with higher [resp., lower] education levels are concentrated in larger [resp., smaller] metropolitan areas. The largest ten metropolitan areas have more than 70% of all college graduates, and 53% of junior high school graduates in their workforce.

Is there any simple rule which holds for the location patterns of industries and of population? For instance, is it possible to infer the industrial composition of a given region from the (population) size and location of the region? Is it possible to explain the variation in the location pattern of workers with different education levels by that of industries, and vice versa? In this paper, we show that the answer to these questions is "Yes", and that in fact the location patterns of industries and of population exhibit strong and persistent regularities. By appropriately identifying the location of industries, a simple relationship is revealed among the number, size and spacing of metropolitan areas in which each industry is located. The regularities in industrial location in turn reveal the location patterns of workers with different education levels. Despite the fact that industries in general have trickled down from larger metropolitan areas to (a larger number of) smaller ones, and the required education level of workers of each industry has decreased over time, the *relative* location of industries (i.e., the relative number, size and spacing of the metropolitan areas where the industries are located) and *relative* worker composition (in terms of education level) of industries have remained unchanged between 1980 and 1995. The stability of the relative location of industries translates likewise into the spatial distribution of population.

There are two key factors in identifying the location of industries. One is the aggregation level of industry and geography. Too much aggregation of industries ends up with grouping industries with different location patterns, and then the characteristics of location (e.g., the population size of a region)

[1]A metropolitan area here is a set of counties which cover a business district and the corresponding residential area. Refer to Section 2 for the precise definition. There were 118 metropolitan areas in 1995.

cannot be a good indicator of its industrial composition.

For the geography, we find that the clear location pattern is detected if "economically cohesive areas" are aggregated. Basically, we find that most industries appear to be agglomerated in metropolitan areas, while each industry is further localized in a small part of a metropolitan area. Apparel companies and publishers, for instance, are often localized in different districts in a metropolitan area, while sharing the same metropolitan advantages in transportation, communication and various business services. That is, the location of industries is in general found in an agglomerated form with a hierarchical nature: the industry-specific agglomeration takes place in a smaller area equivalent to the size of a few counties, and inter-industry one takes place at a larger area equivalent to a metropolitan area in the form of a cluster of neighboring industry-specific agglomerations. Thus, to derive the basic results in this paper, the industry location is identified at the county level, while the property (e.g., the size) of the metropolitan area (rather than the county in which the industry is found) is considered to be relevant for the industry in determining its location in the national landscape. It will be shown that our choice of geographic units reveals distinctively clearer location patterns than other alternatives. Note that when the location of industries exhibits such hierarchical localization, if the specialization of a given area is defined in terms of the employment share of each industry as is often the case (e.g., Henderson [20][21]), the result will be quite different depending on the unit of geographic area (e.g., counties, metropolitan area, prefectures and states). Thus, the choice of the geographic unit is crucial for the characterization of the industry location pattern.

Our definition of regional and industrial categories turned out to be less aggregate than those often used in previous works on the spatial distribution of industries in a country (introduced below). Despite the seemingly complex nature of the geographic and industrial systems using less aggregated units, we will see that the resulting location patterns are (paradoxically) much simpler than in a more aggregated system. It suggests the need for the study of disaggregated systems in both empirical and theoretical analyses of economic location.

The other key factor is the hierarchical nature *among* metropolitan areas: the set of industries found in a smaller metropolitan area is a subset of that found in a larger one. That is, both sporadic and ubiquitous industries are found in large diverse metropolitan areas, but only ubiquitous ones are found in small metropolitan areas. This property is often called *the Christaller's* [3] *hierarchical principle.*[2] The principle is equivalent to what Krugman [28] called *the spatial phase locking* of industrial location. In his work, interactions of multiple industries which are subject to agglomeration economies in various degrees are considered. While all industries geographically concentrate to some extent due to agglomeration economies, their spacing is different depending on the degree of agglomeration economies: the

[2] A similar but less strict hierarchical principle was suggested by Lösch [26].

spacing of industries with a higher degree of agglomeration economies is greater. When demand and/or production externalities exist and are shared by industries, the locations of these industries are synchronized: the location of less ubiquitous industries tends to coincide with that of more ubiquitous ones.[3] Under the hierarchical principle, the rather traditional use of employment share may be inappropriate to characterize the industrial structure in a given geographic area. It likely neglects ubiquitous industries in a large metropolitan area since employment shares of those industries tend to be small compared to the employment shares in a small area. In this paper, the industry-specific agglomeration in a county is identified by its *absolute size*, while a metropolitan area is considered to have all the industries found in at least one of the counties in it. In this way, the hierarchical principle is reflected in the identified location of industries.

There are a bulk of related researches which motivated this paper. Ellison and Glaeser [10] discussed the aggregation problem in the identification of the industrial agglomeration. For the stability of the spatial distribution of population, Black and Henderson [2] for the case of the US, and Eaton and Eckstein [9] for the cases of France and Japan provided evidence in terms of the size distribution of metropolitan areas. For the spatial distribution of industries and its relation to the spatial distribution of population, there are many contributions for the case of the US. To name a few, Black and Henderson [2] and Henderson [20] related the size variation of metropolitan areas to their relative specialization. Duranton and Puga [6] and Henderson [21] proposed measures for the industrial diversity of a metropolitan area, and have shown evidence for positive correlation between the diversity and size of metropolitan areas. Overman and Ioannides [36] investigated the impact of distance between the metropolitan areas on their size and growth. Dobkins and Ioannides [5] and Rauch [38] show evidence for the positive correlation between the size of a metropolitan area and the education level of workers. Black and Henderson [2] compared the specialization of metropolitan areas with the education level of their workers. Though the existing literature suggests the existence of patterns in the industry and population location, the identified patterns are not apparent, and evidence shown is fragmental. In this paper, we propose an alternative approach to identify the location of industries, and relate it to the location of population in a systematic manner. The result is presented for the case of Japan.

The remainder of this paper is organized as follows. The description of the data is given in Section 2. The aggregation problem in the identification of industry location is addressed in Section 3. Upon adjusting the level of aggregation, the location patterns of industries and of population are derived in Sections 4 and 5, respectively. In Section 6, the location of workers with different education levels as well as the worker composition of industries are studied. Finally, we conclude with a discussion on policy and

[3] Informal case studies for the hierarchical principle have been conducted by, e.g., Berry [1], Christaller [3], Dicken and Lloyd [4], Isard [24], Lösch [26], and Marshall [29]. In this paper, we propose a formal test of the principle for the first time.

theoretical implications of our findings.

2 Data

In this paper, we basically look at the location patterns of industries and of population in 1980 and in 1995. The data source and the modification of data we made for the analyses are explained below.

Location and population

The minimum geographical unit we consider is the county. The definition of counties and the county population data are based on the Population Census of Japan in 1980 and 1995.[4] To make the data comparable between the two years, counties in each year are converted to those in 1995. As the location of population and industries, we consider administrative jurisdictions (counties and prefectures), and define the metropolitan area as a set of counties grouped in terms of commuting patterns and population density. More strictly, we follow the definition of the metropolitan area called the *Metropolitan Employment Area* (MEA) developed by Kanemoto and Tokuoka [27], which is comparable to the Core Based Statistical Area (CBSA) of the US.[5] Basically, an MEA consists of a (densely inhabited) business district which has sufficiently large in-flow of commuters and suburb counties in which commuters reside. An MEA may have multiple business districts among which one is the central business district having the positive net inflow of commuters, while the rest are subcenters which have significant commuter flows both to the central business district and from their surrounding counties. We identified 105 MEAs in 1980 and 118 in 1995.[6]

Industries

The employment data we mainly use in this paper are classified according to the three-digit Japanese Standard Industry Classification (JSIC) of the Establishment and Enterprise Census of Japan in 1981 and 1996 which we in turn apply to the 1980- and 1995-data, respectively. Since the industry classification has been subdivided during the 15 years, we aggregate the classification in 1996 to that in 1981. Among the three-digit JSIC-industries, we focus on manufacturing, services, wholesale and retail, which consist of 259 industries. In this paper, we consider these 259 industries as the set of all industries ($\equiv \mathbf{I}$).[7]

Distance

To study the spacing of industries and population, we calculated the distance between a pair of counties as the road-distance between their county halls. The road-distance is based on the network of highways

[4] The county here is equivalent to the *shi-ku-cho-son* in the Japanese Census.
[5] See Office of Management and Budget [34] for the definition of CBSA.
[6] Among 3375 (3370) counties, MEAs include 1329 (1564) counties in 1980 (1995).
[7] From the total of 388 categories in the original classification, heavily regulated industries (tobacco and ordinance) and the "miscellaneous" sectors which do not fit to any of the specific categories are excluded.

and major roads provided by the Japan Map Center (http://www.jmc.or.jp/).[8]

Education

The data on the education level of workers in each county are available only in 1980 and 1990 from the Population Census of Japan. As an approximation, we apply the education level of workers based on the 1990 data to the workers locating in each MEA in 1995 and those employed in each industry in 1996.

3 Identification of industry location

In this section, we compare possible combinations of levels of industry and geographic aggregation in order to find out the best aggregation level to study the industry location pattern on the national landscape. For the industry aggregation, we consider two- and three-digit JSIC codes, while for the geographic aggregation, we have three levels: county, MEA and prefecture. In Section 3.1, as a preliminary analysis, we look at spatial distribution of total employment as well as industry-specific employment, which gives us a clue for determining the appropriate geographic unit to identify the industry location. In Section 3.2, the location pattern of industries revealed under different aggregation levels are compared. In what follows, we provide explanation mainly for 1995, but almost the same aspects are found for 1980 as well unless specified.

3.1 Basic fact of industrial location

In 1995, more than 80% of total employment, and thus, the industrial activities is concentrated in MEAs which account for 30% of the geographical area covered by all counties. Both across and within MEAs, the distribution of industries is far from uniform. This can be verified by comparing the distribution of industry-specific as well as total employment across MEAs with those within each MEA. To do this, it is convenient to use the Theil index in which the inequality in the distribution of workers across all counties can be decomposed into that across and within MEAs.[9] Suppose there are K counties which are grouped into M MEAs; each MEA $m = 1, \ldots, M$ has n_m counties; and out of the total worker population (of a given industry), county $k = 1, \ldots, n_m$ in MEA m has N_{km} workers. Then, we can write the Theil index, T, for the inequality across all counties as follows:

$$T = \frac{1}{K} \sum_{m=1}^{M} \sum_{k=1}^{n_m} (N_{km}/N) \log(N_{km}/N) \qquad (1)$$
$$= T^{inter} + \sum_{m=1}^{M} s_m T_m^{intra},$$

[8]It can be verified that clearer location patterns of industries and population are identified by the road-distance than the simple one-line distance. The reason may be that the road-distance approximates the true distance between locations more precisely than the one-line distance by taking into account geographic obstacles such as mountains, rivers and irregular coastlines.

[9]To calculate the Theil index across counties, we include only the counties that belong to MEAs.

where s_m is the employment share in MEA m, and

$$T^{inter} \equiv \frac{1}{K} \sum_{m=1}^{M} n_m (N_m/N) \log(N_m/N), \tag{2}$$

$$T_m^{intra} \equiv \frac{1}{n_m} \sum_{k=1}^{n_m} (N_{km}/N_m) \log(N_{km}/N_m), \tag{3}$$

where N is the average employment size of all counties, and N_m is that of counties in MEA m. The (cross-county) total Theil index is the sum of the inter-MEA Theil index given by T^{inter} and the weighted average of the intra-MEA Theil index for MEA m ($m = 1, \ldots, M$) given by T_m^{intra}. The value of the index takes 0 when all counties have equal employment size, and increases as employment is concentrated in a fewer number of counties, and takes $\log(K)$ if the employment is completely concentrated in one county.[10] Figure 1 plots for each industry the share of the total Theil index which accrues to the intra-MEA inequality. We can see that for most industries, both inter- and intra-MEA inequality are significant in the total Theil index.

Figure 1

The Theil index for the distribution of workers of all industries is 1.04 (refer to the vertical line in the figure) in which the intra-MEA variation accounts for 81.6% (refer to the horizontal line in the figure).[11] The figure indicates that the spatial distribution of industry-specific employment does not closely follow that of the total employment. For relatively ubiquitous industries (i.e., those with a smaller Theil index), the intra-MEA variation accounts for the most part of the total variation across all counties. It suggests that even the "ubiquitous" industries which are found in most MEAs may be concentrated in a few counties within each MEA. For the relatively sporadic industries (i.e., those with a large Theil index), the share of inter-MEA variation is larger, but the intra-MEA variation still accounts for a significant share. Namely, it is not only that these industries are found in a smaller number of MEAs, but also that they are localized within each of these MEAs. This result suggests that the specialization pattern at the county level and that at the MEA level may appear quite differently.

Next, we have a closer look at the spatial distribution of industries across counties within an MEA. We focus on the (largest) Tokyo MEA as the most illustrative example.[12] On the average, 32.5% of the industry-specific employment is concentrated in three counties among total 229 counties, whereas the largest three counties in total employment account for only 15.4% of total employment in the Tokyo MEA. Figure 2 shows the number of industries for a given share of the top three counties in the industry-specific employment. The graph indicates that most industries exhibit significant agglomeration within the Tokyo

[10] Since the geographic area of counties are not equal, this Theil index is at most a rough measure of spatial concentration.
[11] The Theil index for all the industries is the sum of that for each industry weighted by its share in national employment.
[12] Other MEAs share basically the same property.

MEA. For 90% [resp., 23%] of industries, more than 10% [resp., 50%] of employment is concentrated in the top three (i.e., 1.3%) counties.

Figures 2

An important fact to notice is that in an MEA, industries tend to sort themselves across counties rather than co-locate. Figure 3 depicts frequencies that a county is among the top three counties in the employment of a given number of industries in the Tokyo MEA. There are 125 counties which are among the top three in the employment of at least one industry. Among these counties, 59 (47.2%) are in the top three for only one industry, and 106 (84.8%) for less than ten industries.

Figures 3

We can summarize the results so far as follows:[13]

Fact 1 *(Spatial extent of industrial agglomeration) Industries agglomerate in MEAs, and within each MEA, these industries geographically sort themselves across smaller areas comparable to the size of a few counties. That is, the spatial extent of industry-specific agglomerations is smaller than that of inter-industry ones, and an inter-industry agglomeration can be considered as a cluster of multiple industry-specific agglomerations in close proximity, rather than the agglomeration of the mixture of multiple industries in the same location.*

That is, the location of industries is in general found in an agglomerated form with a hierarchical nature: each industry is localized in a few counties, and several industry-specific agglomerations cluster to form an inter-industry agglomeration which is roughly covered by an MEA. Under such hierarchical agglomeration, the industry location pattern may appear differently depending on the geographic unit chosen to identify the presence of industries. We will address this point in the next subsection.

3.2 Aggregation levels and industry location pattern

By Fact 1, it makes sense to identify the location of industry by the location of industrial agglomeration. Moreover, the agglomeration is hierarchical: as each industry is agglomerated in a small industry district, an inter-industry agglomeration takes place at a greater area consisting of several industry-specific districts. To take this into account, we consider two geographic levels: regions and subregions, so that co-agglomeration of different industries is recognized at the region level, and industry-specific agglomeration at the subregion level. Though the previous subsection implies that the MEA-county pair could be the most appropriate region-subregion pair, we compare other possible alternatives as well in this

[13] Regarding the spatial extent of industrial agglomerations, see Ellison and Glaeser [10, V.D] for a related discussion.

subsection. To identify the *presence* of industry-specific agglomerations in a given subregion, we set the lower threshold for the employment share of the industry in a subregion as follows:

Definition 1 *(Identification of industry-specific agglomeration) Let μ_i and σ_i be mean and standard deviation of the distribution of the employment share (in logarithm) of industry $i \in \mathbf{I}$ in all the areas (at a given geographic aggregation level) in the nation. We say that area ℓ has industry i if and only if $\log s_{\ell i} > \mu_i - \delta\sigma_i$, where $s_{\ell i}$ is the employment share of industry i in area ℓ, and δ is a given positive constant.*[14]

We set $\delta = 1$ for all cases we consider below.[15] The important role of the threshold is to avoid recognizing negligibly small employment as an industrial agglomeration. The screening also excludes subregions which specialize in extremely differentiated products in a given industry, e.g., handmade fountain pen mills in the stationary industry. Here, we examine the employment share instead of employment size because our subregions, e.g., counties, are quite diverse in (population) size. For the most ubiquitous industries such as bakeries and restaurants, the employment size of the industry in a given geographic area is roughly proportional to the total employment size in the area. In general, since large areas employ more than small ones for more ubiquitous industries, these industries in small areas tend to be neglected if we compare across areas the employment size of each industry. To avoid this, we examine employment share, assuming that the employment in ubiquitous industries is approximately proportional to the population size of the subregion. Once the industries located in each subregion are found, the inter-industry agglomeration is identified as follows:

Definition 2 *(Identification of inter-industry agglomeration) Regions are mutually exclusive and non-empty sets of subregions. A region is considered to have all the industries identified in at least one of its subregions as specified in Definition 1.*

We compare five possible region-subregion combinations: (a) county-county, (b) MEA-county, (c) MEA-MEA, (d) prefecture-county and (e) prefecture-prefecture.[16] As a measure of comparison, we use the relationship between the number and average size of regions in which each industry is located.[17] The five combinations of region levels using the three-digit JSIC are compared in Figures 4, while Diagram (b′) shows the MEA-county combination using the two-digit JSIC.[18]

Figures 4

[14] There may be alternative ways to normalize the employment size in each area by using the geographical size of an area, but there seems to be no optimal method to compare the size of industry concentrations in different areas.

[15] The choice of δ does not affect the basic results.

[16] Note that an MEA may extend over the prefecture boundaries.

[17] This is one of the relationships which characterize the industry location pattern we will be discussing in the following sections. The result of comparison among different aggregation levels is basically the same in all the measures of industrial location pattern that will be introduced.

[18] The 388 three-digit industries are aggregated into 51 two-digit industries.

Regardless the geographic aggregation, clearer patterns emerge in the three-digit industry classification than in the two-digit one (though only the MEA-county case is shown in the figure for the two-digit case). Namely, as is clear from the comparison between Diagrams (b) and (b′), there are much larger cross-industry variations in both the number and average size of industry location using the three-digit than the two-digit classification. By increasing aggregation to the two-digit JSIC level, information on the variety of size and number of industry locations is lost significantly, and these two variables are no longer good measures for distinguishing industries. Thus, in order to understand the location pattern of industries, the three-digit classification is a better aggregation level.

Next, we turn to the comparison among spatial aggregation levels focusing on the three-digit industry classification. A remarkably clear pattern can be recognized in the MEA-county case (type b), which can be characterized by the log-linear relation between the average size ($\overline{\text{SIZE}}$) and the number (#MEA) of MEAs in which a given industry is located. The ordinary least square estimation (OLS) gives the following result:[19]

$$\log(\overline{\text{SIZE}}) = \underset{(0.005514)}{7.404}{}^{**} - \underset{(0.002883)}{0.7080}{}^{**} \log(\#\text{MEA}), \quad R^2 = 0.9958. \tag{4}$$

In Section 3.1, we learned that the spatial extent of an MEA [resp., a county] roughly matches that of inter-industry [resp., industry-specific] agglomeration. Moreover, in Section 4, we will see that the location of a given industry appears to be strongly influenced by the presence of co-located industries, and that the size of an MEA indeed reflects the type of industries located there. Thus, if there is a simple relation between the number and size of locations for each industry, it is not surprising that the MEA-county case reveals it most precisely. For the county-county case (type a), there is a larger variance (toward the smaller value) in the average county size for the industry locating in a smaller number of counties. Here, the size of a county in which a given industry is located does not convey the information on other co-located industries, since the spatial extent of inter-industry agglomeration is greater than that of a county. Thus, the size of the counties cannot correctly characterize the location of industries. For the MEA-MEA case (type c), there are several industries whose average MEA size is significantly smaller than that calculated in the MEA-county case. This is due to the fact that an MEA contains multiple industry-specific agglomerations, and that a larger MEA contains agglomerations of a wider variety of industries (to be subsequently confirmed in Section 5.1). This causes a problem in the identification of industry-specific agglomerations. Namely, when the employment shares at the MEA level are compared, the industry agglomeration in a larger MEA tends to be less represented than that in a smaller MEA. Thus, the industries which are identified in both large and small MEAs in the MEA-county case tend to be erroneously found only in small MEAs in the MEA-MEA case. For the prefecture-county case (type d), the fit is relatively good, though it is not as good as the MEA-county

[19]* significant at 5% level, ** significant at 1% level. Standard errors are in parentheses.

case. However, if the industry-specific agglomeration is identified at the prefecture level, the clear pattern does not appear (type e). Again, this result suggests that the county is the most appropriate among the available geographical units to identify industry-specific agglomerations. As far as the industry-specific agglomerations are identified at the county level, prefectures may be roughly considered as appropriate geographic units where inter-industry agglomerations take place.

Based on the comparison above, in the following sections, we derive all the results under the MEA-county combination of the geographic aggregation level with the three-digit industry classification. For convenience, we call an MEA in which a given industry is located *an industry-choice MEA*.

4 Location patterns of industries

In this section, we study the relationship among the four key variables of industry location: the (i) *number*, (ii) *average (population) size*, (iii) *spacing of industry-choice MEAs*, and (iv) the *critical (MEA population) size* of each industry. The first three are self-explanatory. The critical size of an industry (to be defined in Section 4.2) is the threshold value of the size of an MEA such that the industry in question is likely to be located in the MEA with size above the threshold level. We show that the location pattern of industries identified in terms of these key variables has been clear and persistent over the 1980-1995 period, despite the fact that the value of each key variable significantly changed in the same period.

4.1 Industry location and the population size of an MEA

Figure 3(b) in the last section depicts the clear relation between the number and average size of industry-choice MEAs for year 1995.[20] The corresponding OLS result is shown in eq. (4). The relation in 1980 is equally clear ($R^2 = 0.9921$). Pooling the data, we estimate the log-linear models with time (year) dummy variables. The F-tests among the models with and without dummies indicate that both the intercept and slope are significantly different (at 1% level) between the two years. However, the difference in slope is very small (i.e., less than 1% of the slope estimated from the samples in either year). The change in the intercept reflects that in the number of industry-choice MEAs for each industry which increased by 16.0% on average. Note that the slope represents the elasticity of the average industry-choice MEA size with respect to the number of industry-choice MEAs. It follows that the relative location of industries (in terms of the elasticity between the number and size of industry-choice MEAs) has remained unchanged, while each industry has become more ubiquitous. We will revisit the inter-temporal change in Section 4.3.

Next, let us see the spacing of industry which is defined as follows:

[20]The list of industries ordered by the number of locating MEAs in 1995 is in Appendix.

Definition 3 *(Industry spacing) Let* $\mathbf{L}$ *be the set of all MEAs in which a given industry is located. For each* $m \in \mathbf{L}$*, define the set of neighboring agglomerations of the same industry by* $\mathbf{N}_m \equiv \{i | i \in \mathbf{L} - \{m\}$ *and* $\|m, i\| < \max(\|m, j\|, \|i, j\|)\ \forall j \in \mathbf{L} - \{m, i\}\}$*. The spacing of the industry is given by* $\frac{1}{|\mathbf{L}|} \sum_{m \in \mathbf{L}} \frac{1}{|\mathbf{N}_m|} \sum_{i \in \mathbf{N}_m} \|m, i\|$.[21,22]

Under this definition, $\mathbf{N}_m$ includes only the closest MEA among those located in a *similar direction* from MEA m. To illustrate this, suppose that the spacing is measured in terms of one-line distance instead of road-distance. The geographic relationship among MEAs $i, j \in \mathbf{N}_m$ and MEA m cannot look like Diagram (a) in Figure 5, but it should look like Diagram (b). In Diagram (a), MEA j cannot be a member of $\mathbf{N}_m$, since there is another MEA i which has the specified industry and is located between MEA m and j. In Diagram (b), though MEA j is located closer to MEA m than MEA i, it is on the opposite side of MEA i with respect to MEA m. Thus, both MEAs i and j can be considered as the direct neighbors of MEA m. In the one-dimensional location space, $\mathbf{N}_m$ includes at most two MEAs (i.e., only the direct neighbors).

Figure 5

Figure 6 plots the industry spacing (I_SPACING) versus the number of industry-choice MEAs in logarithm for year 1995.[23]

Figure 6

The corresponding OLS result is given as

$$\log(\text{I\_SPACING}) = \underset{(0.01662)}{2.677}^{**} - \underset{(0.008685)}{0.4575}^{**} \log(\#\text{MEA}), \quad R^2 = 0.9152. \tag{5}$$

The fit is good for year 1980 as well ($R^2 = 0.8452$), and both intercept and slope are not significantly different between the two years. Again, we see the persistent relative location of industries (in terms of the elasticity between the industry spacing and the number of industry-choice MEAs). It is to be noted that the industry spacing has decreased by 7.7% on average between 1980 and 1995, which is consistent with the fact that industries are becoming ubiquitous (recall the increase in the number of industry-choice MEAs). Thus, the plot in the figure has shifted right and below which happened to end up with insignificant difference in the intercepts (not only the slopes) in 1980 and 1995.

Thus, the relationship among the number, average size and spacing of industry-choice MEAs is characterized by a simple persistent pattern over the 15 years.

[21] By the location of an MEA, we mean the location of the county hall of the central county of the MEA.

[22] Under this definition, $\mathbf{N}_m$ includes only one MEA among those located in a similar direction from MEA m. For instance, in the one-dimensional location space, $\mathbf{N}_m$ includes at most two MEAs (i.e., only the direct neighbors).

[23] The spacing for each industry can be found in the list of industries in Appendix.

4.2 Hierarchical principle of industrial location

It has been pointed out that industrial composition of cities is hierarchical since the seminal work by Christaller[3]. Namely, the so-called Christaller's hierarchical principle holds: the set of industries found in a less diverse city are the subset of that found in a more diverse city. The principle has had little attention in the modern urban and regional economics, and has never been seriously tested.[24] However, if it is true, the principle should have an important implication for the identification of industrial structure of a city. In this section, we conduct a simple test for the hierarchical principle in the industrial composition of MEAs.

Figure 7 demonstrates the principle observed in the actual location pattern of industries. Each point in the figure corresponds to one realized industry location in 1995 such that the industry is characterized by the number of industry-choice MEAs (vertical axis), and the location (i.e., MEA) by its industrial diversity (horizontal axis).

Figure 7

Notice that the plots are sparse near the southwest corner, meaning that the industries with a smaller number of locations are found only in MEAs with large industrial diversity. On the other hand, MEAs with small industrial diversity have only ubiquitous industries (i.e., those locating in a large number of MEAs). It follows that the list of industries found in a less diverse MEA is mostly contained in that found in a more diverse MEA.

It is true that given a finite and discrete number of locations and the variation in the number of locations across industries, such a hierarchical structure would likely appear even under random location of industries. As a random location of industries which on average results in the same industrial diversity of each MEA in reality, we consider the following situation. Suppose we place boxes (=MEAs) of different sizes (=industrial diversity) in a row, and prepare balls (=industry agglomerations) with different colors (=different industries) where the numbers of balls vary depending on the color. (The number of balls for each color is smaller than the number of boxes, as the number of industry-choice MEAs is smaller than the total number of MEAs.) We throw balls toward boxes. There is a trick in the boxes such that each box can have at most one ball for each color; otherwise, the probability that each ball drops in a given box is proportional to the size of the box. In this setup, we will have the industrial diversity of MEAs in reality on average under the random location of industries with given number of industry-choice MEAs.

To test the randomness of industry location (under the given number and industrial diversity of MEAs) in the above mentioned context, we apply the binary Logit model to estimate the probability

[24] The large body of the theoretical and empirical studies for the specialization of cities is based on the model by Henderson [19][20] whose framework implies that each city is essentially specialized in a single industry.

that a given industry is located in a given MEA:

$$\Pr(\text{located=true}) = \frac{\exp[F(\text{DIV}, \#\text{MEA})]}{1 + \exp[F(\text{DIV}, \#\text{MEA})]}, \tag{6}$$

where F is a function of the industrial diversity of the MEA (DIV) and the number of industry-choice MEAs. The following two specifications of F are compared:[25]

Full model :

$$F(\text{DIV}, \#\text{MEA}) = \underset{(0.3875)}{-13.306^{**}} + \underset{(0.001902)}{0.04983^{**}}\text{DIV} + \underset{(0.004407)}{0.08752^{**}}\#\text{MEA} - \underset{(2.26E-05)}{0.00132^{**}}\text{DIV} \times \#\text{MEA} \tag{7}$$

Restricted model :

$$F(\text{DIV}, \#\text{MEA}) = \underset{(0.1743)}{-11.306^{**}} + \underset{(0.000744)}{0.03956^{**}}\text{DIV} + \underset{(0.000742)}{0.06224^{**}}\#\text{MEA} \tag{8}$$

If the industry location is random, we expect that the location probability increases with DIV as well as with #MEA *independently*. That is, the cross-effect of DIV×#MEA should not increase the explanatory power of the model. The log-likelihood ratio between the two models is 34.74, which indicates a significant difference (at 1% level) on the goodness of fit between the two models. That is, the location choice of more ubiquitous industries and that of less ubiquitous industries are affected differently by the industrial diversity of each location (i.e., MEA). Hence, the randomness of the industry location is rejected. The coefficients of DIV and #MEA are positive and significant in either specification as expected. The negative coefficient of DIV×#MEA in the full model implies that the industrial diversity (positively) matters more for the location of industries with a smaller number of industry-choice MEAs.[26]

The above result is a statistical confirmation of the hierarchical principle. Following Christaller [3], we call an industry which locates in a smaller [resp. larger] number of MEAs a *higher-order* [resp., *lower-order*] industry, and call a more [resp., less] diversified MEA a *higher-order* [resp., *lower-order*] *MEA*.[27] The principle is equivalent to the spatial phase locking of industrial location suggested by Krugman [28]. Namely, industry location follows a seemingly synchronized pattern. Why do industrial clusters attract each other? The existing literature suggests three typical mechanisms that lead to the spatial phase locking. Namely, due to direct and indirect production linkages, concentration of different industries creates thick markets for their input and output (e.g., Fujita [12]; Fujita and Hamaguchi [15]; Helseley and

[25]Coefficients are estimated for the samples in year 1995 (the number of samples=30562).

[26]Due to the possible multicolinearity, we cannot be fully confident regarding the value of each coefficient. However, the basic result holds even if samples are restricted to manuracturing, services, or wholesale and retail.

[27]Higher-order industries are typically either subject to large scale economies (e.g., coke, briquette, blast furnace manufacturing, petroleum refining) or highly specialized (e.g., fur, leather, surveying equipment, fireproof product, spectacle manufacturing, special school education services, social and cultural science research services). Lower-order industries are subject to high tranport cost in the general sense. Examples of these are the manufacture and wholesale/retail of perishable products (e.g., meat and dairy food, vegetable and fruit food), that of heavy product (e.g., stone and related product, cement), and services/retails of frequent use (e.g., attorney services, department strores, automobile maintenance, drug and cosmetic retail). See Appendix for the detail.

Strange [18]), and induces innovations for production technologies (Duranton and Puga [6]). Similarly, an agglomeration of industries means a large market for the final good and services because of the concentration of workers, which in turn attracts even larger agglomeration of final good producers (e.g., Fujita, Krugman and Venables [16, Ch.11]). Another important reason for the co-agglomeration of industries is the sharing of urban infrastructure such as transport hubs. Concentration of industries and population promotes improvement of urban infrastructure, and vice versa (e.g., Fujita, Krugman and Venables [16, Ch.13]; Konishi [25]; Mori and Nishikimi [31]).

Under the hierarchical principle, it is possible to infer which industries are located or sustainable in a given MEA once the diversity of the MEA is known. Under the *strict* hierarchical principle, there is a *critical (MEA) diversity* for each industry such that these industries locate in MEAs with diversity greater than the critical level. As we will see in Section 5.1, the industrial diversity and population size of an MEA have a strong positive correlation. Thus, we can also define *the critical (MEA) size* for each industry in the same manner. For the stochastic counterpart of the hierarchical principle, we can define the critical diversity/size in the following way:

Definition 4 *(Critical diversity and size of an industry) The critical diversity/size of an industry is the threshold level of diversity/size of an MEA estimated by the binary Logit model such that the probability that an MEA has a given industry is greater [resp., smaller] than 0.5 if the diversity/size of the MEA is greater [resp., smaller] than the critical level.*[28]

To save space, we only present the result for the critical size. We first estimated by the maximum likelihood method the logit model in which the location probability of each industry is determined by MEA size. The likelihood ratio test and the Hosmer-Limeshow test exhibit that the model fits well for most industries.[29] Figure 8 plots the estimated critical size (C_SIZE) in logarithm versus the number of industry-choice MEAs for each industry for year 1995.

Figure 8

The corresponding OLS result is given as

$$\log(\text{C\_SIZE}) = \underset{(0.03466)}{6.60}^{**} - \underset{(0.000409)}{0.01692}^{**}(\#\text{MEA}), \quad R^2 = 0.9181. \tag{9}$$

[28]More precisely, the critical diversity of each industry is estimated by the equation: Pr(located=true) = $\exp(\alpha + \beta\text{DIV})/[1 + \exp(\alpha + \beta\text{DIV})]$, and is given by $-\alpha^*/\beta^*$, where α^* and β^* are estimated values of α and β. For the critical size, DIV is replaced by the log(MEA size). It is to be noted that the location probability which corresponds to the critical diversity/size does not have to be 0.5 depending on the objective.

[29]Of the total 259 industries, the critical size is well defined for 189 industries (e.g., industries locating in all MEAs are excluded). For 14 industries out of the 189, the null hypothesis that the coefficient of MEA size equals zero is not rejected at 10 percent level by the likelihood ratio test. For 19 out of the 189, the Holsmer-Limeshow test rejects the hypothesis that the discrepancy between expected and observed responses is zero. In Figure 7 and the estimation of eq.(9), we excluded these industries.

Only the intercept is significantly different between the two years. The critical size decreased by 1.26% on average. We summarize our finding in this subsection as follows:

Fact 2 *(Hierarchical principle) Industries located in a smaller MEA are likely to be found in a larger MEA more than random location would imply. In particular, industries locating in fewer MEAs tend to be found in larger MEAs. The estimated ratio of the critical size of an industry locating in* n *MEAs to that of an industry locating in* m *MEAs is estimated as* $\exp(-0.017[n-m])$*, where* n *and* m *are positive integers, and remained unchanged over the 1980-1995 period.*

Some remarks are in order regarding the relation between the hierarchical principle of industrial location and specialization of MEAs. In the urban economics literature, a popular view is that cities are more or less specialized in a single or few industries in order to utilize the industry specific *localization economies* (Black and Henderson [2, Sec.3] and Henderson [20][21]). In particular, Henderson [21, p.590] claims that roughly half of the 243 USA metropolitan areas are highly specialized in one particular (the three-digit SIC) industry. Is his finding inconsistent with ours, i.e., with the hierarchical principle? The answer is "No." It is true that some industries are subject to strong localization economies, and tend to dominate in the medium or small cities. In Japan, there are eight MEAs which have employment share greater than 10% in a single industry.[30] However, the hierarchical principle implies that MEAs relatively specialize in higher-order industries. Figure 9 shows this point by plotting the employment share versus the relative order of industries in each MEA for year 1995.[31]

Figure 9

The plot clearly indicates that a higher-order industry tends to have a higher employment share in an MEA. In particular, the mean employment share is above the national share only for the industries whose intra-MEA order is in the top 6%. That is, the specialization of MEAs is a natural outcome of the hierarchical principle. Also, Henderson [21] points out that medium and small cities have zero employment in most industries as an indication of prevalence of specialized metropolitan areas. But, this is another natural outcome of the hierarchical principle. These relatively small metropolitan areas tend to have only a small set of industries, while the large ones have almost all industries.

Finally, though it is popular to view the location of industries in terms of relative specialization of metropolitan areas, it overlooks the industrial diversity of large metropolitan areas under the hierarchical

[30] Three MEAs specialize in motor vehicles, parts and accessories, three in electronic parts and devices, and one in household electric appliances, and one in seafood products.

[31] The horizontal axis indicates the industry's order in a given MEA in terms of the number of industry-choice MEAs. The order is higher (i.e., is of smaller value) if the number of industry-choice MEAs is smaller, and the lowest order (i.e., the largest value) is normalized to 1. The vertical bar indicates mean $\pm$ one standard deviation of employment shares within an MEA of the industries at a given order relative to the industry's national employment share.

principle.[32,33] A larger metropolitan area not only specializes in higher-order industries but also have significant employment in lower-order industries, while a smaller one has only lower-order industries. As a result, the regularities in industrial location pattern comparable to what we have found in this paper have not been reported in the existing literature based on the relative specialization. Thus, in revealing industry location patterns, the choice of aggregation levels is just as important as the choice of methods for identifying industries in a given area.

4.3 Inter-temporal change in industry location

Between 1980 and 1995, the average increases in the number and average size of industry-choice MEAs are 16.0% and 2.8%, respectively, while the industry spacing decreased by 7.7%. In the meantime, the number of MEAs has increased by 12.3%, while the increase in the MEA size is on average 13.1%, and not surprisingly, there is no significant change in the distance between the neighboring MEAs. Notice that the increase in the number of industry-choice MEAs is larger [resp., smaller] than that in the number [resp., the average size] of MEAs. We can say that industries are on average becoming more ubiquitous, and that they are gradually spreading out from larger MEAs into smaller ones. It is consistent with the change in the critical size which decreased by 1.26% on average. This decentralization can be verified by looking at the change in location pattern of each industry. Figure 10 compares the value in 1995 and that in 1980 (in logarithm) of each key variable for each industry. Table 1 shows the corresponding OLS result assuming a log-linear relation.

Figure 10 and Table 1

The fits are good for all the cases. The log-measured slope and intercept are both significantly smaller than 1 (at 1% level) for every key variable. This result is consistent with the fact that industries are becoming more ubiquitous. The number of industry-choice MEAs became relatively smaller for the industries that were already ubiquitous in 1980, since obviously these industries will be able to increase the number of locations at most at the rate of increase in the number of MEAs (12.3%). The average MEA size has become relatively smaller for the industries with average MEA size larger in 1980. Recall that the industries are decentralizing from larger MEAs to smaller ones. But, since the number of MEAs are limited, industries that were already fully ubiquitous in 1980 cannot decrease the average MEA size anymore. As a result, an industry with originally larger average MEA size tend to exhibit a greater

[32]See, for instance, Figure 5(c) which plots the average MEA size of industries when the presense of each industry is identified at the MEA level rather than at the county level. Several industries have much smaller average MEA size compared to the case of Diagram (b) where the industry presence is identified at the county level. In Diagram (c), these industries are not identified in large MEAs which has many industries.

[33]An example of the measure of industrial concentration in terms of relative specialization can be found in Ellison and Glaeser [10] who identify the concentration of an industry by the concentration of firms in the industry relative to the concentration of employment of all industries in a given region. For another example, see Black and Henderson [2, Sec.3] and Henderson [20, Sec.1.2] which effectively group metropolitan areas in the US in terms of relative specialization.

decrease in the average MEA size. A similar explanation can be given for the change in industry spacing. The greater decrease in the critical size for the industries whose critical size was larger in 1980 further confirms the tendency that the existing industries are becoming more ubiquitous over time.[34]

Although the location pattern of each industry is changing over time, compared to the intertemporal changes in the values of the key variables for each industry, the relative location pattern of industries exhibits surprisingly persistent regularities as we have seen in the previous subsections. We can say the following:

Fact 3 *(Industry location pattern) There is a strong and persistent negative correlation between the number of industry-choice MEAs and each of the average industry-choice-MEA size, spacing and critical size of industries. In particular, the elasticity between each pair of the number, average size and spacing of industry-choice MEAs has been constant over the 1980-1995 period, while the industries on average have become more ubiquitous, and in particular, they have decentralized from larger MEAs to smaller ones.*

Although the "existing" industries are becoming more ubiquitous, it is also true that the simultaneous formation of new industries is not captured in our data defined according to 1980 industry classification. Notable examples are the computer industries in the 80s and information technology (IT) such as internet-related industries in the 90s. These new industries are often found in large cities. In the year 2000, among the software, information processing and internet related (i.e., IT) industries found in the Yellow Page, 46.7%, 10.7% and 4.5% are located in the largest three MEAs: Tokyo, Osaka and Nagoya, respectively.[35] The mechanism that leads to the continuous formation of new industries at the largest cities together with the decentralization of the existing industries creates the clear and persistent industrial location patterns characterized by Facts 2 and 3.

5 Industrial location and spatial distribution of population

In this section, we look at the spatial distribution of population, i.e., the size and location of MEAs, and relate it to the location pattern of industries we have found so far. Here, we focus on the three key variables: (i) the *industrial diversity* (i.e., the number of locating industries), (ii) *size*, and (iii) *spacing* of an MEA.

5.1 Industrial diversity, size and spacing of MEAs

Figure 11 shows the relation between the size (SIZE) and industrial diversity of MEAs in 1995. It is apparent that the two variables have strong positive correlation. Although it appears that the largest three MEAs (Tokyo, Osaka and Nagoya) have almost all industries, it may be due to the available industry

[34] See Henderson, Kuncoro and Turner [22] for evidence of the decentralization in the case of the US.
[35] Data source: Ministry of Land, Infrastructure and Transport of Japan (http://www.mlit.go.jp/).

classification. Namely, the three-digit JSIC may not be appropriate in distinguishing the industrial composition of these largest MEAs[36]

Figure 11

Taking this limitation into account, we conduct regression analysis excluding the largest three MEAs. The OLS result is given by

$$\log(\text{DIV}) = \underset{(0.04643)}{1.516}^{**} + \underset{(0.008305)}{0.1374}^{**}\log(\text{SIZE}), \quad R^2 = 0.7025. \tag{10}$$

There is no significant difference in the estimated coefficients between the two years.[37]

Next, we look at the relation between the size and spacing of MEAs. Recall that by the hierarchical principle (Fact 2), an MEA with size greater than the critical size of a given industry would likely have industries which are of higher-order than that industry. Taking this into account, here, , we are concerned with the distance between a given MEA and its neighboring MEAs with *equal* or *larger* size, i.e., MEAs of equal or higher-order.[38] Thus, we define the spacing of MEAs (analogous to that of industry spacing) as follows:

Definition 5 *(Spacing of MEAs) Let* $\mathbf{M}(S)$ *be the set of all MEAs whose population size is greater or equal to* S*. Define the set of neighboring MEAs in* $\mathbf{M}(S)$ *of MEA* $m \in \mathbf{M}(S)$ *by* $\mathbf{N}_m(S) \equiv \{i | i \in \mathbf{M}(S) - \{m\}$ *and* $\|m, i\| < \max(\|m, j\|, \|i, j\|) \forall j \in \mathbf{N}_m(S) - \{m, i\}\}$*. The spacing of MEAs with population size at least* S *is given by* $\frac{1}{|\mathbf{M}(S)|}\sum_{m\in\mathbf{M}(S)} \frac{1}{|\mathbf{N}_m(S)|}\sum_{i\in\mathbf{N}_m(S)} \|m, i\|$..

Figure 12 plots the spacing of MEAs (M-SPACING) for the size greater than or equal to the size of each MEA in logarithm for year 1995.

Figure 12

The values of the spacing corresponding to the population size of the seven smallest MEAs in the figure are almost equal, since it is basically the avarage spacing of all MEAs. Taking this into account, we

[36]For instance, the formation of the new IT industries discussed in Section 4.3 which are found mostly in Tokyo is not reflected in this industry classification.

[37]The correlation between the industrial diversity and size of MEAs appears more clearly compared to those reported in the existing literature. For instance, as a measure of industrial diversity of metropolitan areas, the Herschman-Herfindahl index (HHI) is often used (e.g., Henderson [21]), which is defined as the sum of the squared employment shares for each industry in a given metropolitan area. A possible problem in using this measure is that it is influenced by the difference in the labor requirement across industries. Duranton and Puga [6] measures the diversity of a metropolitan area by the *relative diversity index* (RDI) defined by the inverse of the sum of the absolute difference between the employment share of each industry in a metropolitan area and that in the nation. Thus, unlike the HHI, the RDI takes into account the variation in the labor requirement across industries. However, the index takes the maximum value when the employment shares in a metropolitan area coincide with those of the nation, the meaning of which is not clear. As a result, in their studies, the correlation between the size and industrial diversity of MEAs appear to be weak. Under our data set, the replacement of our measure of industrial diversity of an MEA by HHI and RDI in eq.(10) results in $R^2 = 0.075$ and 0.510, respectively.

[38]Our preliminary study indicated that the proximity to the smaller (i.e., lower-order) MEAs has little influence on the value of other key variables. See Overman and Ioannides [36] for a related discussion.

conduct regression analysis excluding these samples. The OLS result is given by[39]

$$\log(\text{M\_SPACING}) = \underset{(0.05265)}{-0.7669}^{**} + \underset{(0.009356)}{0.4845}^{**}\log(\text{SIZE}), \quad R^2 = 0.9654. \tag{11}$$

There is no significant difference in the estimated coefficients between 1980 and 1995.

It is to be noted that given the hierarchical principle of industrial structure of MEAs, the relationship between the spacing and average size ($\overline{\text{SIZE}}$) of MEAs with size greater than or equal to a given value should roughly reproduce the regularity revealed between the industry spacing and average size of industry-choice MEAs (refer to Fact 3; Figure 4(b) and Figure 6). The following OLS result confirms this point:

$$\log(\text{M\_SPACING}) = \underset{(0.07613)}{-1.675}^{**} + \underset{(0.01209)}{0.5748}^{**}\log(\overline{\text{SIZE}}), \quad R^2 = 0.9780. \tag{12}$$

Indeed, no significant difference in the estimated coefficients can be found if we instead regress industry spacing against average size of industry-choice MEAs.

5.2 Intertemporal change in the industrial diversity, size and spacing of MEAs

Finally, we summarize the changes in the values of the key variables for the MEAs that existed in 1980 and 1995, and conclude with some stylized facts. Figure 13 compares the value in 1995 to that in 1980 of each key variable, and Table 2 shows the corresponding OLS results.

Figure 13 and Table 2

The number of locating industries in MEAs increased by 5.9% on average. Since the industrial composition of MEAs is hierarchical (Fact 2), and since we use the same industry classification for both years, the average increase in the number of locating industries indicates that the MEAs with originally small diversity are diversifying over time. This can also be confirmed by Diagram (a) and the regression result in Table 2(row 2). In particular, the log-measured slope is 0.81 and is significantly smaller than 1 (at 1% level). This result is consistent with Fact 3 that industries are becoming more ubiquitous.[40]

Furthermore, we compare the employment share of each industry in each MEA in 1995 and that in 1980. By pooling all the 259 industries and 102 MEAs which are common in the two years, the OLS result assuming a log-linear relation between the employment shares in the two years is given in row (5) of the table. The log-measured slope is 0.80 which is significantly smaller than 1 (at 1% level). Given the hierarchical industrial composition of MEAs, we can roughly say that industries with larger employment

[39] Here, only the MEAs that are connected by roads are included. In particular, 13 MEAs in two islands, Hokkaido and Okinawa, are excluded from the samples.

[40] The diversification of large metropolitan areas may be underestimated, since the emergence of new industries which mostly happens in large metropolitan areas is not taken into account in our fixed industry classification. See the related discussion in Section 4.3.

shares are those in the small MEAs. The OLS result implies that these industries have experienced greater decrease in the employment share, which is consistent with our finding that industries are becoming more ubiquitous. Namely, industries that were located only in large MEAs before are now found in smaller ones as well.

The population size of an MEA increased by 13.1% on average. However, the relative size of MEAs is remarkably stable between 1980 and 1995 (Figure 13(b) and Table 2(row 3)): the log-measured slope is not significantly different from 1 with $R^2 = 0.99$.[41] There is rather large change in the population size of a few MEAs. For instance, the population sizes of Fukuoka, Kanazawa and Utsunomiya have grown by 24.6%, 75% and 56.9%, respectively. At the same time, however, the MEAs located near these growing MEAs experienced a relatively small increase or even a decrease in size: e.g., the changes in population size of Boufu, the population sizes of Nagasaki and Tokuyama near Fukuoka were 3.4%, 2.2% and -8.5%, respectively, while those of Ashikaga, Kiryu and Takasaki near Utsunomiya were -0.1%, -7.8% and 14.8%, respectively. Thus, a growth in size of an MEA is associated with a decline in size of other surrounding MEAs. As a result, the size distribution of MEAs have stayed nearly unchanged.

The spacing of MEAs decreased by 2.86% on average reflecting the increase in the number of MEAs. The relative spacing of MEAs has been stable between 1980 and 1995 (Figure 13(c) and Table 2(row 4)): the log-measured slope is not significantly different from 1 with $R^2 = 0.96$. While the relative size and spacing of MEAs are stable over the studied period, recall that the change in the location pattern of each industry, and thus, the specialization of each MEA, is significant as shown in Section 4.3. Hence, we can say the following:

Fact 4 *(Industrial diversity, size and location of MEAs) There is a strong and persistent positive correlation between the industrial diversity and size, and between the spacing and size of MEAs. In particular, the relative value of the key variables as well as the relative size and relative spacing of MEAs have remained unchanged over the 1980-1995 period, despite a significant increase in the industrial diversity and in the size of MEAs during the same period.*

6 Education level and location of workers

We have seen that workers are geographically concentrated in MEAs. However, depending on the education level, the degree of concentration is not the same. Figure 14 shows the Lorenz curves of the distribution of workers across MEAs in 1990. The Lorenz curves are depicted separately for the total worker population, and for workers with each of the four education levels, junior high school, high school, community/technical college and college.[42]

[41] The stability of the city size distribution has been reported by many authors for different nations. See, e.g., Black and Henderson [2] for the US, Eaton and Eckstein [9] for Japan and France.

[42] The horizontal axis ranks MEAs in terms of their shares of workers with a given education level (the rank is normalized by the total number of MEAs, 118). The vertical axis reports the cumulated share of workers of a given education level in

Figure 14

The figure clearly indicates that workers with higher education levels are more geographically concentrated.[43] In this section, we characterize MEAs and industries by the education level of their workers.

6.1 Education level of workers in MEAs

The data show that workers with higher education are more easily found in larger MEAs. Figure 15 plots in logarithm the share of workers at each education level in the total employment in each MEA in 1995. The composition of workers with each education level exhibits a visible pattern. Namely, while the share of high school graduates is similar across MEAs, that of a higher education level is larger in a larger MEA, and that of lower education level is larger in a smaller MEA.[44] The corresponding OLS results are shown in Table 3 (columns 2-4).

Figure 15 and Table 3 (columns 2-4)

Though the goodness of fit is generally low ($R^2 < 0.16$), given the clear pattern of industry and population location (Sections 4 and 5), this may be partly due to the variation in labor requirement (in both quantity and education level) across industries. Namely, the education level of workers in a more labor intensive industry is more pronounced in the worker composition of a given MEA. In fact, the relative worker composition of MEAs is remarkably stable. Figure 16 plots the value (in logarithm) in 1995 versus that in 1980 of the share of workers at each education level in each MEA. The corresponding OLS results are given in Table 3 (columns 5-7). The fits are good for each education level ($R^2 > 0.96$).

Figure 16 and Table 3 (columns 5-7)

It is clear from the figure that education level of workers has increased on average. On the other hand, in all the education levels, the slopes are not significantly different from 1. What this implies is that though the education level of workers are in general increasing over time, the *relative* spatial distribution of workers with different education levels has remained unchanged. That is, the growth rate of the share of workers with higher education levels is higher in MEAs which originally had a larger share of them. We summarize our finding so far as follows:

Fact 5 *(Education level of workers in MEAs) There is a weak positive correlation between the population size of an MEA and the average education level of workers in an MEA. The relative worker composition*

the MEAs up to a given position in the rank.

[43]The value of the Gini index for the distribution of the total population is 0.69, while the corresponding value for the cases of workers with education up to junior high school, high school, community/technical college, and college, is 0.60, 0.69, 0.77, and 0.82 respectively.

[44]Dobkins and Ioannides [5] obtained a similar result for the U.S.

in terms of the education level among MEAs has remained the same, while the education level of workers has increased on average over the 1980-1990 period.

6.2 Education level of workers for industries

The available data do not reveal the education level of workers in each industry. Thus, we infer it from the education level of workers in industry-choice MEAs. Namely, as a substitute for the share of workers with a given education level in a given industry, we calculate the weighted average of the share in industry-choice MEAs (where the weight is the MEA size). The result for 1995 is shown in Figure 17, where the industries are distinguished in terms of the number of industry-choice MEAs along the horizontal axis. The OLS results assuming a log-linear relation are given in Table 4 (columns 2-4). The fits are good except for high school graduates ($R^2 > 0.98$). The slopes are not significantly different between the two years except for the case of high school graduates. Note that the share of high school graduates are similar across industries (Figure 17). Thus, we can conclude that the relative worker compositions in the industries have essentially remained unchanged in the studied period.

Figure 17 and Table 4 (columns 2-4)

Figure 18 plots the employment share (in logarithm) of workers at each education level in 1995 versus that in 1980 for each industry. The corresponding OLS results are given in Table 4 (columns 5-7). From the figure, it is apparent that the education level of workers in each industry has become higher. The table indicates a systematic change in the worker composition of each industry: the log-measured slope for the college graduate share is 0.86 (row 5), and is significantly smaller than 1 (at 1% level), while that for the junior high school graduate share is 1.3 (row 2), and is significantly larger than 1. It follows that industries which are originally with a larger share of college graduates experienced relatively smaller increase in the share of college graduates, and relatively smaller decrease in the share of junior high school graduates.

Figure 18 and Table 4 (columns 5-7)

We can say that as industries become ubiquitous (Fact 3), their demand for the workers with higher education level decreases. It may be due to the fact that the production technologies became more standardized over time, so that the education level of workers becomes irrelevant for productivity.[45] It is surprising, though, that the relative worker composition of industries has remained unchanged, despite the decentralization and the change in the worker composition of each industry. We summarize our finding as follows:

[45] See Duranton and Puga [7] for a micro-foundation relating the standardization in the production technology and the decentralization of industries.

Fact 6 *(Education level of workers in industries) There is a strong and persistent negative correlation between the average education level of workers for industries and the number of industry-choice MEAs. In particular, the elasticity between the number of industry-choice MEAs and the share of workers at each education level for industries has been constant during the 1980-1995 period, while the average education level of workers has increased. From the view point of each industry, the average education level of workers has decreased comparably more in the industries that had originally a larger share of highly educated workers as they have become more ubiquitous.*

It is to be noted that workers are sorted in terms of education level across industries and locations. On one hand, if the difference in the industry-specific technologies and the market conditions induces variation in the location pattern of industries, then the location determinant from the view point of firms may be responsible for the resulting variation in the spatial distribution of workers according to their education level. On the other hand, the difference in the productivity between the workers depending on their education level may induce a different choice of residential location, and accordingly firms choose their location seeking for the workers with appropriate education levels (Mori and Turrini [33]).

7 Discussions

In this paper, we proposed a new approach to identify the location pattern of industry and population. The structure of the spatial economy detected has important implications for the regional development policies as well as theoretical development of spatial economics. We address each issue in the subsections below. We close discussion with possible extensions of our research.

7.1 On regional development policies

Which industry is sustainable in a given MEA? To answer this question, our analysis suggests that one needs to look at both the MEA specific factors and the structure of the global regional system. The former is usually well taken, and it is important for adding information on the idiosyncratic characteristics of the location which may be in favor of particular industries.[46] However, often, the regional industrial policies neglect the latter. For instance, recently in Japan, several periphery cities try to attract new IT industries to boost their economy, simply because these industries are the fastest growing ones. But, we know that these industries are the most sustainable in large MEAs. Our finding indicates that there is not much degree of freedom in the location pattern of industries: there is a stable relationship between the number, size and spacing of the MEAs in which a given industry can be located. For the size, in particular, the critical size provides a useful benchmark. Namely, to promote the development of a given industry in a given MEA, the critical size of the industry should not be much greater than the size of

[46] An example is the so-called first nature advantage of location such as the presence of natural harbors and oil fields.

the MEA in question. The requirement of education level of workers for the industry further narrows down the possibility. Consideration of the global pattern of industry location will help identify feasible industries in a given MEA, and should enhance the efficiency of regional industrial development.

It is also important to view the spatial distribution of population via an appropriate geographical unit. The population of a given region is often considered as an indicator of the prosperity of the region. The recent empirical studies for the Japanese economic geography have consensus on the on-going monopolarization of population in Japan towards Tokyo, and relative decline in the rest of the nation since 80s.[47] However, their result crucially depends on the choice of geographic units (usually utilizing administrative jurisdictions) on which the spatial distribution of population is defined. As we have seen in this paper, the choice of geographic units should reflect the spatial extent of the actual population concentration. If the population distribution is viewed through MEAs, then such monopolarization is not evident, and the relative size of MEAs is very stable instead.

7.2 Theoretical perspective

Most theoretical results on economic location are based on simple “highly aggregated” models. One of the popular simplifications of the location space is the two-region setup. In this setting, for instance, numerous models of economic agglomerations have been developed.[48] These are important in formalizing the mechanism of geographic concentration which is a general tendency of industry location as we have seen in Section 3.1. However, when it comes to explaining the actual spatial distribution of industries on a more general location space such as the national landscape, it is not very useful. They are of little help for regional industrial development policies: they hardly provide ways to quantitatively evaluate the possibility of the development of a given industry in a given city/region. In particular, in these simplified models, depending on the parameter values, drastically different spatial configurations become equilibria (e.g., complete concentration of industries in one region or complete spread of industries across regions). As shown in our results, in reality, there is not much degree of freedom in the location pattern of industry. That is, the self-organization of the spatial economy leads to the formation of simple location patterns of industries, despite the possibility of numerous alternative locations for each industry. The basic reason why the simplified model is so helpless in explaining the actual industry location pattern we believe is too much aggregation for both geography and industry. Our results suggest the need for the theoretical development with more disaggregated models. It is to be noted that increasing the number of regions does not improve the situation if the spatial relationship among the regions are neglected. A notable example is Henderson [19] which allows the number of cities to be endogenous, while there is no inter-city

[47] See, e.g., Fujita and Tabuchi [13].
[48] See, e.g., Fujita and Thisse [14] for a survey.

spatial structure.[49] His work provided an important benchmark at the early stage of theoretical (and also empirical) development in explaining the relation between the size and specialization of cities. However, our finding suggests that *the variation in the distance between cities* turned out to be one of the key element to understand the self-organization in industry and population location.

Finally, we learned that the spatial distribution of workers is different depending on their education level. In particular, the workers with higher education level tends to geographically more concentrate, and the relative distribution of workers with each education level among MEAs has been stable. Moreover, the location of workers is linked with that of industries in a simple manner. Our result suggests the need for another disaggregation. Namely, the presence of heterogeneous (in the level of human capital) workforce may be a crucial factor that generates the difference in the location pattern of industries, and also for the interregional inequalities.[50] Again, in most theoretical models, there is no consideration of such variety in workforce as a source of difference in the resulting location pattern of industries.[51]

7.3 Extension

Here, we present two possible extensions of our research. First, in our analysis, we find stable location patterns of industry and population on the national location space in Japan. But, do such pattern formations take place only at the national level? In other words, what is the geographic area in which the self-organization of economic activities generates certain patterns? Can we observe the same regularities in smaller subregions of Japan[52], and also in different nations? It is an interesting question to be asked in future research.

Second, we distinguished industries by their location pattern. However, it is true that there may be a wide variation in the location pattern of organizational units (e.g., headquarters, research and development, manufacturing plant, etc.) of a given firm depending on their role in the firm. In fact, it has been pointed out that the location pattern may also be distinguished among the intra-firm organizations, and that there is a positive correlation between the population size of a city and the number of cooperate control linkages generated in the city (e.g., Fujita and Tabuchi [13]; Pred [37]; Ross [39]). It is not at all obvious how the intra-firm spatial organization relates to what we have found in this paper on the industry location pattern. However, given the rather generic presence of multi-unit firms, it is an important agenda for future research.[53]

[49] Another example is Tabuchi, Thisse and Zheng [40] which allow the endogenous number of regions, but each pair of regions are assumed to be *equi*-distant.

[50] See, e.g., Fingleton [11], Fujita and Tabuchi [13], and Glaeser, Scheinkman and Shleifer [17] for the evidence of interregional income inequality in the EU, Japan and the US, respectively.

[51] Exception are Monfort and Ottaviano [30] and Mori and Turrini [33] which derived the result consistent with our finding on the spatial distribution of industries with explicit consideration of worker heterogeneity in terms of skill level.

[52] See Mori and Nishikimi [32] for the preliminary analysis.

[53] See, e.g., Duranton and Puga [8] and Ohta and Fujita [35] for the recent theoretical development for the location of multi-unit firms.

References

[1] Berry, B., *Geography of Market Centers and Retail Distribution*, Prentice Hall (1968).

[2] Black, D. and Henderson, V., "Urban evolution in the USA," mimeograph, Brown University (1998).

[3] Christaller, W., *Die Zentralen Orte in Suddeutschland*, Gustav Fischer (1933), English translation by C.W. Baskin, *Central Places in Southern Germany*, Prentice Hall (1966).

[4] Dicken, P. and Lloyd, P.E., *Location in Space: Theoretical Perspectives in Economic Geography*, 3rd ed., New York: HarperCollins (1990).

[5] Dobkins, L. and Ioannides, Y., "Dynamic evolution of the size distribution of U.S. cities," in J.M. Huriot and J.-F. Thisse (eds.), *Economics of Cities*, Cambridge: Cambridg University Press, 217-260 (2000).

[6] Duranton, G. and Puga, D., "Diversity and specialization in cities: why, where and when does it matter?" *Urban Studies 37, 535-555 (2000).*

[7] Duranton, G. and Puga, D., "Nursery cities: urban diversity, process innovation, and the life-cycle of product," *American Economic Review*, forthcoming.

[8] Duranton, G. and Puga, D., "From sectoral to functional urban specialization," Mimeograph, Department of Geography and Environment, London School of Economics (2001).

[9] Eaton, J. and Eckstein, Z., "Cities and growth: theory and evidence from France and Japan," *Regional Science and Urban Economics* 27, 443-474 (1997).

[10] Ellison, G. and Glaeser, E.L., "Geographic concentration in U.S. manufacturing industries: a dartboard approach," *Journal of Political Economy* 105, 889-927 (1997).

[11] Fingleton, B., "Specification and testing of Markov chain models: an application to convergence in the European Union," *Oxford Bulletin of Economics and Statistics* 59, 385-403 (1997).

[12] Fujita, M., "Location of firms with input transactions," *Environmental and Planning A* 13, 1401-1414 (1981).

[13] Fujita, M. and Tabuchi, T., "Regional growth in postwar Japan," *Regional Science and Urban Economics* 27, 643-670 (1997).

[14] Fujita, M. and Thisse, J.-F., *Economics of Agglomeration*, forthcoming, Cambridge: Cambridge University Press, Ch.2 (2001).

[15] Fujita, M. and Hamaguchi, N., "Intermediate goods and the spatial structure of an economy," *Regional Science and Urban Economics* 31, 79-109 (2001).

[16] Fujita, M., Krugman, P. and Venables, A., *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press (1999).

[17] Glaeser, E., Scheinkman, J.A. and Shleifer, A., "Economic growth in a cross-section of cities," *Journal of Monetary Economics* 36, 117-143 (1995).

[18] Helseley, R. and Strange, W., "Matching and agglomeration economies in a system of cities," *Regional Science and Urban Economics 20, 189-222 (1990)*

[19] Henderson, J.V., "The sizes and types of cities," *American Economic Review* 64, 640-656 (1974).

[20] Henderson, V., *Urban Development*, Oxford: Oxford University Press (1988).

[21] Henderson, V., "Medium size cities," *Regional Science and Urban Economics* 27, 583-612 (1997).

[22] Henderson, V., Kuncoro, A. and Turner, M., "Industrial development in cities," *Journal of Political Economy* 103, 1067-1090 (1995).

[23] Ioannides, Y.M. and Overman, H.G., "Spatial evolution of the US urban system," Discussion paper No.0018, Tufts University (2000).

[24] Isard, W., *Location and Space-Economy*, Cambridge, MA: MIT Press (1956).

[25] Konishi, H., "Hub cities: formation of a hub without increasing returns," *Journal of Urban Economics* 48, 1-28 (2000).

[26] Lösch, A., *The Economics of Location*, Jena, Germany: Fischer (1940), English translation, New Haven, CT: Yale University Press (1954).

[27] Kanemoto, Y. and Tokuoka, K., "The proposal for the standard definition of the metropolitan area in Japan," Discussion paper No.J-55, Center for International Research on the Japanese Economy, Tokyo University, Tokyo, Japan (2001)

[28] Krugman, P., *The Self-Organizaing Economy*, Cambridge: Blackwell (1996).

[29] Marshall, J., *The Structure of Urban Systems*, Toronto, University of Toronto Press (1989).

[30] Monfort, P. and G.I.P. Ottaviano, "Local labor markets, skill accumulation and regional disparities," Mimeograph, Department of Economics, Bocconi University (2000).

[31] Mori, T. and Nishikimi, K., "Economies of transport density and industrial agglomeration," *Regional Science and Urban Economics*, forthcoming.

[32] Mori T. and Nishikimi, K., "Recursive pattern formation in the industry and population location," (in progress) Institute of Economic Research, Kyoto University, Japan (2001).

[33] Mori, T. and Turrini, A., "Skills, agglomeration and segmentation," CEPR Discussion paper No.2645 (2000).

[34] Office of Management and Budget, "Standards for defining metropolitan and micropolitan statistical areas," Federal Register Vol.65, No.249 (2000).

[35] Ohta, M. and Fujita, M., "Communication technology and spatial configurations of intrafirm units and residential units," *Journal of Regional Science* 21, 87-104 (1991).

[36] Overman, H. and Ioannides, Y.M., "Cross-sectional evolution of the US city size distribution," Mimeograph, Department of Geography and Environment, London School of Economics (2000).

[37] Pred, A., *The Spatial Dynamics of U.S. Urban-Industrial Growth*, Cambridge: MIT Press (1966).

[38] Rauch, J., "Productivity gains from geographic concentration of human capital: evidence from the cities," *Journal of Urban Economics* 34, 380-400 (1993).

[39] Ross, S., *The Urban System and Networks of Corporate Control*, Greenwich, CT: JAI Press (1991).

[40] Tabuchi, T., Thisse, J.-F. and Zheng, D.-Z., "On the evolution of multi-regional system," Discussion paper 2001/3, CORE, Université catholique de Louvain, Belgium (2000).

Appendix. List of industries (ordered by the number of industry-choice MEAs)

Order	ID	Name	#MEA	Spacing (km)
1	131	Leather glove manufacturing	9	179 (4)
2	134	Fur manufacturing	10	151 (6)
3	119	Coke manufacturing	11	180 (3)
3	369	Racing facility services	11	234 (1)
5	120	Briquette manufacturing	12	154 (5)
5	145	Blast furnace iron manufacturing	12	123 (12)
7	128	Industrial leather product manufacturing, excluding gloves	14	149 (7)
7	427	Special school education services	14	197 (2)
9	127	Leather manufacturing	18	127 (11)
9	146	Non-blast furnace iron manufacturing	18	143 (9)
11	98	Pulp manufacturing	20	131 (10)
12	86	Fur product manufacturing	21	144 (8)
13	193	Scientific equipment manufacturing	23	119 (13)
14	195	Spectacle manufacturing and lenses polishing	24	118 (15)
15	423	Primary education services	25	89 (27)
16	117	Petroleum refining	26	101 (17)
16	191	Surveying equipment manufacturing	26	91 (24)
18	188	Aircraft and parts manufacturing	28	119 (14)
19	74	Silk manufacturing	30	109 (16)
19	118	Lubricant manufacturing, excluding petroleum refining by-products	30	69 (59)
19	123	Tire and tube manufacturing	30	88 (29)
22	66	Sugar manufacturing	33	90 (26)
23	153	Non-ferrous metal refining, primary	34	100 (18)
24	129	Leather footgear material manufacturing	35	88 (30)
25	140	Fireproof product manufacturing	38	68 (63)
25	440	Social and cultural science research services	38	99 (20)
27	141	Carbon and graphite product manufacturing	39	74 (47)
28	113	Chemical fiber manufacturing	40	77 (40)
28	186	Bicycle and parts manufacturing	40	70 (57)
28	196	Watch, clock and parts manufacturing	40	93 (22)
28	370	Racing	40	90 (25)
28	437	Rehabilitation services	40	94 (21)
33	4417	Slaughtering services	41	93 (23)
34	206	Musical instrument and record manufacturing	42	89 (28)
35	142	Abrasive product manufacturing	44	68 (67)
36	75	Yarn mill product manufacturing	45	81 (33)
37	185	Railroad equipment and parts manufacturing	46	81 (34)
38	81	Lace and special textile goods manufacturing	48	63 (95)
39	155	Non-ferrous metal and alloy rolling	50	71 (53)
40	147	Steel and steel product manufacturing	51	72 (51)
40	367	Theatre services, excluding movie projection	51	87 (31)
42	211	Lacquer ware manufacturing	52	77 (39)
43	424	Secondary education services	53	69 (62)
44	124	Rubber and plastic footgear manufacturing	54	74 (46)
44	130	Leather footgear manufacturing	54	76 (41)
44	132	Bag and case manufacturing	54	78 (38)
44	375	Public broadcasting services	54	99 (19)
48	76	Thread manufacturing	56	64 (89)
49	167	Boiler, engine and turbine manufacturing	58	73 (50)
49	443	Scientific and cultural organizaiton services	58	78 (37)
51	71	Oil and fat manufacturing	59	68 (65)
52	382	Blacksmith services	60	76 (42)
53	149	Plated iron material manufacturing	61	67 (71)
53	194	Optical equipment manufacturing	61	73 (49)
53	348	Cheap lodging services	61	79 (35)
56	110	Chemical fertilizer manufacturing	62	84 (32)
57	80	Rope and netting manufacturing	63	76 (43)
57	343	Office machinery rental services	63	75 (45)
59	154	Non-ferrous metal refining, secondary, including non-ferrous alloy manufacturing	64	69 (61)
60	99	Paper manufacturing	65	71 (54)
60	138	Construction clay product manufacturing, excluding ceramics	65	63 (91)
62	101	Paper product manufacturing	66	62 (100)
62	133	Handbag and small case manufacturing	66	73 (48)
62	178	Lighting equipment manufacturing	66	67 (70)
62	368	Performance services	66	75 (44)
66	77	Fabric manufacturing	67	64 (84)
67	157	Wire and cable manufacturing	68	68 (69)
68	376	Commercial broadcasting services	70	78 (36)
69	114	Chemically processed oil product and paint manufacturing n.e.c.	71	55 (166)
69	171	Textile machinery manufacturing	71	65 (76)
69	181	Electrical measuring equipment manufacturing	71	67 (72)
72	95	Religious equipment manufacturing	72	67 (73)
72	208	Office and artist's instrument manufacturing	72	60 (110)
74	209	Accessories, buttons and related goods manufacturing, excluding precious metal and jewelry products	73	62 (102)
75	159	Tin-plated product manufacturing	74	64 (88)
75	361	Custody and lease services	74	60 (112)
77	112	Organic chemical product manufacturing	75	63 (96)
77	205	Precious metal product manufacturing, including jewel processing	75	62 (98)
77	408	Maternity clinic services	75	66 (75)
80	168	Agricultural machinery manufacturing, excluding agricultural equipments	76	62 (101)
81	402	Literary and artistic professional services	77	65 (78)
82	414	Health consultation services	78	70 (55)
82	444	Political organization services	78	68 (64)
84	381	Furniture repair services	79	69 (60)
85	177	Household electric equipment manufacturing	80	64 (83)
85	230	Agents and brokerage	80	68 (66)
85	349	Boarding services	80	70 (56)

The orders in terms of industry spacing (in decending order) are in parentheses.

88	108	Bookbinding and related services	81	61	(104)
88	148	Steel product manufacturing n.e.c., excluding plated iron materials	81	68	(68)
88	432	Social insurance services	81	72	(52)
91	389	Inquiry services	82	70	(58)
92	68	Non-alcoholic beverage manufacturing	84	64	(85)
92	125	Industrial rubber product manufacturing	84	54	(182)
92	156	Non-ferrous cast product manufacturing	84	58	(130)
92	187	Ship and equipment manufacturing	84	65	(80)
96	426	University education services	85	61	(105)
96	439	Natural science research services	85	66	(74)
98	115	Drug and medicine manufacturing	86	64	(82)
99	84	Underwear manufacturing , excluding Japanese style	87	63	(92)
99	214	Textile product wholesale, excluding clothing and apparel	87	59	(126)
99	393	Mass certificaiton	87	64	(87)
102	105	Publishing, excluding newspaper	88	64	(81)
102	190	Measuring equipment manufacturing	88	59	(121)
102	213	Miscellaneous merchandise wholesale	88	65	(79)
105	341	Miscellaneous goods rental service	89	64	(86)
106	180	Electronic equipment manufacturing	90	61	(107)
106	192	Medical equipment manufacturing	90	63	(93)
108	139	Ceramic product manufacturing	91	60	(117)
108	161	Heating apparatus and plumbing supply manufacturing	91	56	(148)
108	174	Office and household machinery and equipment manufacturing	91	56	(154)
108	179	Communication equipment manufacturing	91	58	(131)
108	392	Merchandise inspecting services	91	61	(106)
113	111	Inorganic chemical product manufacturing	92	63	(94)
113	164	Metal wire product manufacturing n.e.c.	92	60	(111)
113	165	Bolt, nut and rivet manufacturing	92	60	(116)
113	364	Movie production and distribution	92	65	(77)
117	79	Textile dyeing and finishing	93	58	(136)
117	100	Processed paper manufacturing	93	56	(159)
119	169	Construction and mining machinery manufacturing, including construction, agricultural and transport tractors	94	62	(97)
120	121	Paving material manufacturing	95	61	(109)
120	350	Special lodging	95	63	(90)
122	207	Playing and sporting goods manufacturing	96	59	(124)
122	357	Special bathing services	96	62	(99)
122	446	Auditorium services	96	59	(119)
125	78	Knit fabric manufacturing	98	58	(133)
125	107	Plate making for printing	98	59	(122)
125	372	Park services	98	59	(120)
128	136	Glass and glass product manufacturing	99	58	(135)
129	65	Grain mill product manufacturing	100	57	(145)
129	69	Alcoholic beverage manufacturing	100	57	(139)
129	70	Feed and organic fertilizer manufacturing	100	58	(128)
129	383	"Hyogu"* services *Papering, mounting and related services	100	56	(156)
129	425	High school education services	100	57	(140)
134	91	Wooden container manufacturing	101	60	(113)
134	430	Social education services	101	59	(125)
136	172	Special-industrial machinery and equipment manufacturing	102	57	(138)
137	104	Newspaper publishing	103	61	(103)
137	353	Dyeing and related services	103	59	(123)
137	365	Movie projection	103	60	(114)
140	170	Metal processing machinery manufacturing	104	57	(142)
140	356	Ordinary bathing services	104	61	(108)
140	401	Commercial and engineering designing services	104	58	(129)
143	62	Seafood manufacturing	105	60	(115)
143	150	Forged and cast steel manufacturing	105	54	(181)
143	182	Electronic and communication parts manufacturing	105	57	(144)
143	184	Automobile and parts manufacturing	105	54	(203)
143	242	Dried food retail	105	60	(118)
148	83	Outerwear manufacturing, excluding Japanese style	106	57	(143)
148	163	Metal processing, excluding enamelling	106	54	(255)
148	391	Secretarial and mimeographic services	106	58	(132)
148	397	Attorney services	106	57	(141)
152	160	Tableware, tool and ordinary hardware manufacturing	107	56	(161)
152	388	News services	107	58	(127)
154	64	Condiment manufacturing	108	54	(180)
154	90	Wooden fabricated material manufacturing	108	58	(137)
154	176	Distributive and industrial electric machinery manufacturing	108	55	(162)
157	173	Ordinary industrial machinery and equipment manufacturing	109	54	(187)
157	442	Labour organization services	109	58	(134)
159	61	Meat and dairy food manufacturing	110	56	(149)
159	143	Stone and related product manufacturing	110	55	(163)
159	250	Japanese restaurants n.e.c.	110	54	(194)
162	63	Vegetable and fruit food manufacturing	111	56	(153)
162	377	Wired sound sevices	111	56	(157)
162	395	Private employment agent services	111	56	(150)
165	89	Wooden material manufacturing, excluding furniture	112	56	(158)
165	215	Clothing and apparel wholesale	112	54	(204)
165	218	Meat wholesale	112	53	(258)
165	220	Other agricultural product wholesale	112	55	(168)
165	239	Alcoholic beverage and condiment retail	112	55	(172)
165	390	Advertising services	112	55	(164)
165	421	Christian services	112	56	(151)
165	428	Kindergarten services	112	56	(147)
165	435	Aged welfare services	112	57	(146)
174	344	Automobile rental services	113	56	(152)
174	387	Information services	113	53	(259)
174	411	Dental mechanic services	113	55	(169)
174	417	Industrial waste management services	113	56	(155)
174	419	Shinto services	113	56	(160)
179	102	Paper container manufacturing	114	55	(165)
179	420	Buddhist services	114	54	(188)
179	441	Business organization services	114	55	(171)
182	137	Cement and cement product manufacturing	115	54	(186)

182	231	Department stores	115	54	(185)
182	267	Second-hand goods retail n.e.c.	115	55	(174)
182	342	Production machinery and equipment services	115	54	(198)
182	379	Parking services	115	55	(170)
182	429	Miscellaneous education services	115	55	(167)
182	434	Children's welfare services	115	55	(179)
182	436	Handicapped welfare services	115	54	(184)
190	94	Furniture manufacturing	116	53	(256)
190	216	Cereal wholesale	116	55	(175)
190	217	Vegetable and fruit wholesale	116	54	(183)
190	219	Seafood wholesale	116	55	(177)
190	227	Furniture and household equipment wholesale	116	53	(256)
190	256	Bicycle retail, including motorcycles	116	54	(195)
190	258	Hard-and kitchen wave retail, excluding agricultrual equipment	116	55	(173)
190	259	Glass and ceramics retail	116	55	(178)
190	269	Camera and photographic materials retail	116	54	(202)
190	398	Notary and scrivenery services	116	55	(176)
200	210	Plastic product manufacturing n.e.c.	117	54	(205)
200	223	Chemical product wholesale	117	54	(195)
200	224	Mineral and metal material wholesale	117	54	(205)
200	241	Seafood retail	117	54	(188)
200	243	Vegetable and fruit retail	117	54	(205)
200	245	Cereal retail	117	54	(188)
200	248	Noodle shops	117	54	(197)
200	257	Furniture and fixture retail	117	54	(205)
200	347	Hotel services	117	54	(188)
200	354	Barbering services	117	54	(188)
200	358	Linen supply	117	54	(199)
200	360	Clothing repairing and related services	117	54	(201)
200	362	Funeral services	117	54	(188)
200	371	Sporting facility services	117	54	(200)
200	399	Accounting and auditing services	117	54	(205)
215	67	Bakery and confectionery product manufacturing	118	54	(205)
215	96	"Tagegu*"manufacturing *Slides and screens in Japanese style	118	54	(205)
215	106	Printing, excluding mimeographing	118	54	(205)
215	162	Construction metal product manufacturing, including cannery sheet metal manufacturing	118	54	(205)
215	221	Food and beverage wholesale	118	54	(205)
215	222	Drug and cosmetic wholesale	118	54	(205)

215	225	Machinery and equipment wholesale	118	54	(205)
215	226	Building material wholesale	118	54	(205)
215	228	Scrap material wholesale	118	54	(205)
215	233	"Kimono", cloths and bedding retail	118	54	(205)
215	234	Men's dress retail	118	54	(205)
215	235	Ladies' and children's dress retail n.e.c.	118	54	(205)
215	236	Footgear retail	118	54	(205)
215	238	General food retail	118	54	(205)
215	240	Meat retail	118	54	(205)
215	244	Confectioneries and bakeries	118	54	(205)
215	247	Restaurant n.e.c.	118	54	(205)
215	249	"Sushi" shops	118	54	(205)
215	251	Bars and cabarets	118	54	(205)
215	252	"Sake" and beer halls	118	54	(205)
215	253	Tea rooms	118	54	(205)
215	255	Motor vehicle retail	118	54	(205)
215	260	Household appliance retail	118	54	(205)
215	262	Florist	118	54	(205)
215	263	Drug and cosmetic retail	118	54	(205)
215	264	Agricultural articles retail	118	54	(205)
215	265	Fuel retail	118	54	(205)
215	266	Book and stationery retail	118	54	(205)
215	268	Sports, toys, musical instruments and other recreation goods retail	118	54	(205)
215	270	Optician	118	54	(205)
215	346	Recreation goods rental services	118	54	(205)
215	352	Laundry services	118	54	(205)
215	355	Beautifying services	118	54	(205)
215	359	Photographic services	118	54	(205)
215	373	Recreation hall services	118	54	(205)
215	378	Automobile maintenance	118	54	(205)
215	380	Machinery repair services	118	54	(205)
215	394	Building services	118	54	(205)
215	400	Construction services	118	54	(205)
215	403	Tutoring services	118	54	(205)
215	405	Hospital services	118	54	(205)
215	406	Ordinary clinic services n.e.c.	118	54	(205)
215	407	Dental clinic services	118	54	(205)
215	409	Pseudo-medical services	118	54	(205)
215	416	Domestic waste management services	118	54	(205)

(1) variables	(2) intercept	(3) value in 1980	(4) R2
(2) number of industry-choice MEAs	0.17857** (0.040489)	0.93859** (0.021930)	0.8757
(3) average MEA size	0.33659** (0.13912)	0.94561** (0.022973)	0.8674
(4) spacing	0.1103 (0.05952)	0.9200** (0.03219)	0.7607
(5) critical size	0.4059 (0.3426)	0.9057** (0.06415)	0.5910

Table 1. Inter-temporal change in the key variables for industries

(1) variables	(2) intercept	(3) value in 1980	(4) R2
(2) industrial diversity	0.4649** (0.0619)	0.8057** (0.02726)	0.8973
(3) population size	-0.00991 (0.08585)	1.010** (0.01532)	0.9875
(4) spacing	0.07277 (0.04177)	1.030** (0.02106)	0.9641
(5) employment share of each industry	-1.2239** (0.034278)	0.79715** (0.004608)	0.5615

Table 2. Inter-temporal changes in the characteristics of MEAs

(1) share in 1995	(2) intercept	(3) SIZE in 1995	(4) R2	(5) intercept	(6) share in 1980	(7) R2
(2) junior high school	-0.44342** (0.006206)	-9E-09** (1.97E-09)	0.1521	0.059567 (0.10014)	0.98630** (0.019966)	0.9606
(3) high school	-0.33080** (0.0029880)	1.58E-10 (9.46E-10)	0.000242	0.20752** (0.077364)	0.98140** (0.015401)	0.9760
(4) community/technical college	-1.1172** (0.007555)	8.72E-09** (2.39E-09)	0.1027	0.16497** (0.072012)	1.0161** (0.017424)	0.9714
(5) college	-1.0908** (0.009899)	1.46E-08** (3.13E-09)	0.1584	0.19616** (0.058892)	1.0024** (0.014031)	0.9808

Table 3. Worker composition of MEAs

(1) share in 1995	(2) intercept	(3) #MEA	(4) R2	(5) intercept	(6) slare in 1980	(7) R2
(2) junior high school	-0.6883** (0.001038)	0.08221** (0.000714)	0.9810	0.04249** (0.00806)	1.304** (0.006363)	0.9939
(3) high school	-0.33879** (0.00050)	0.004852** (0.000261)	0.5728	-0.40571** (0.0.007851)	-0.2162** (0.02230)	0.2679
(4) community/technical college	-0.8836** (0.001060)	-0.06810** (0.000554)	0.9833	0.1666** (0.004407)	1.0017** (0.003746)	0.9964
(5) college	-0.6794** (0.001557)	-0.1174** (0.000814)	0.9878	-0.02748** (0.002556)	0.8619** (0.002517)	0.9978

Table 4. Worker composition of industries

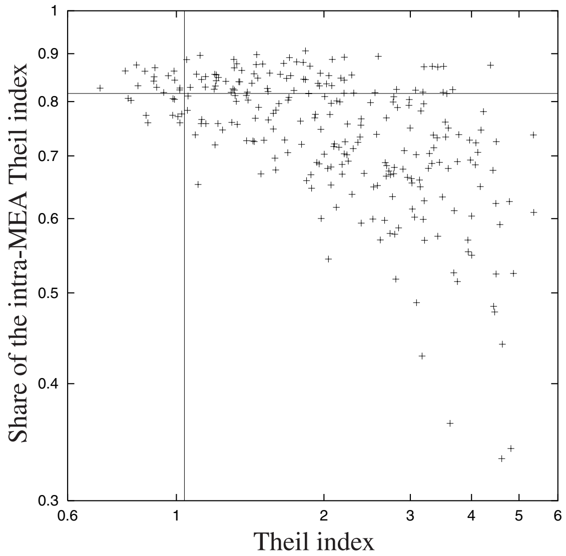


Figure 1. Lumpy industrial location pattern across and within MEAs

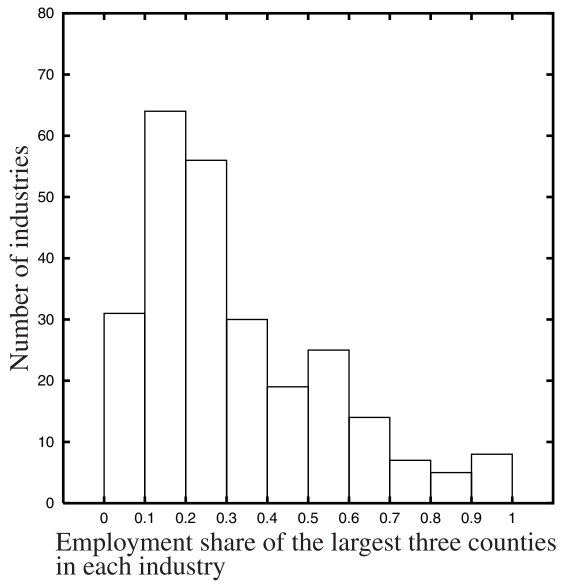


Figure 2. Industry-specific localization within the Tokyo MEA

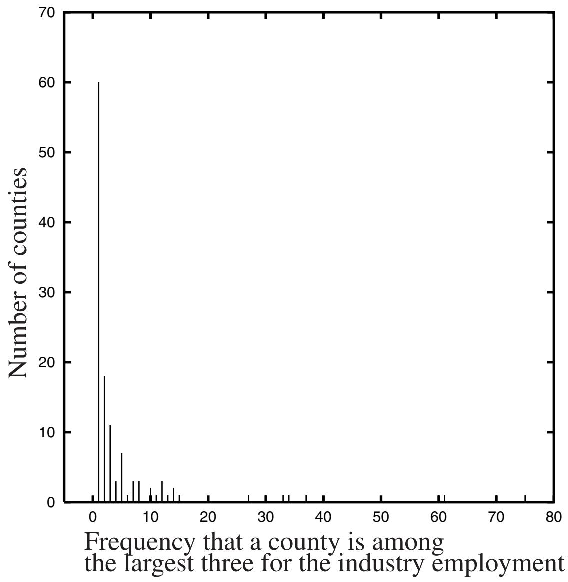


Figure 3. Location variety of industry-specific localization within the Tokyo MEA

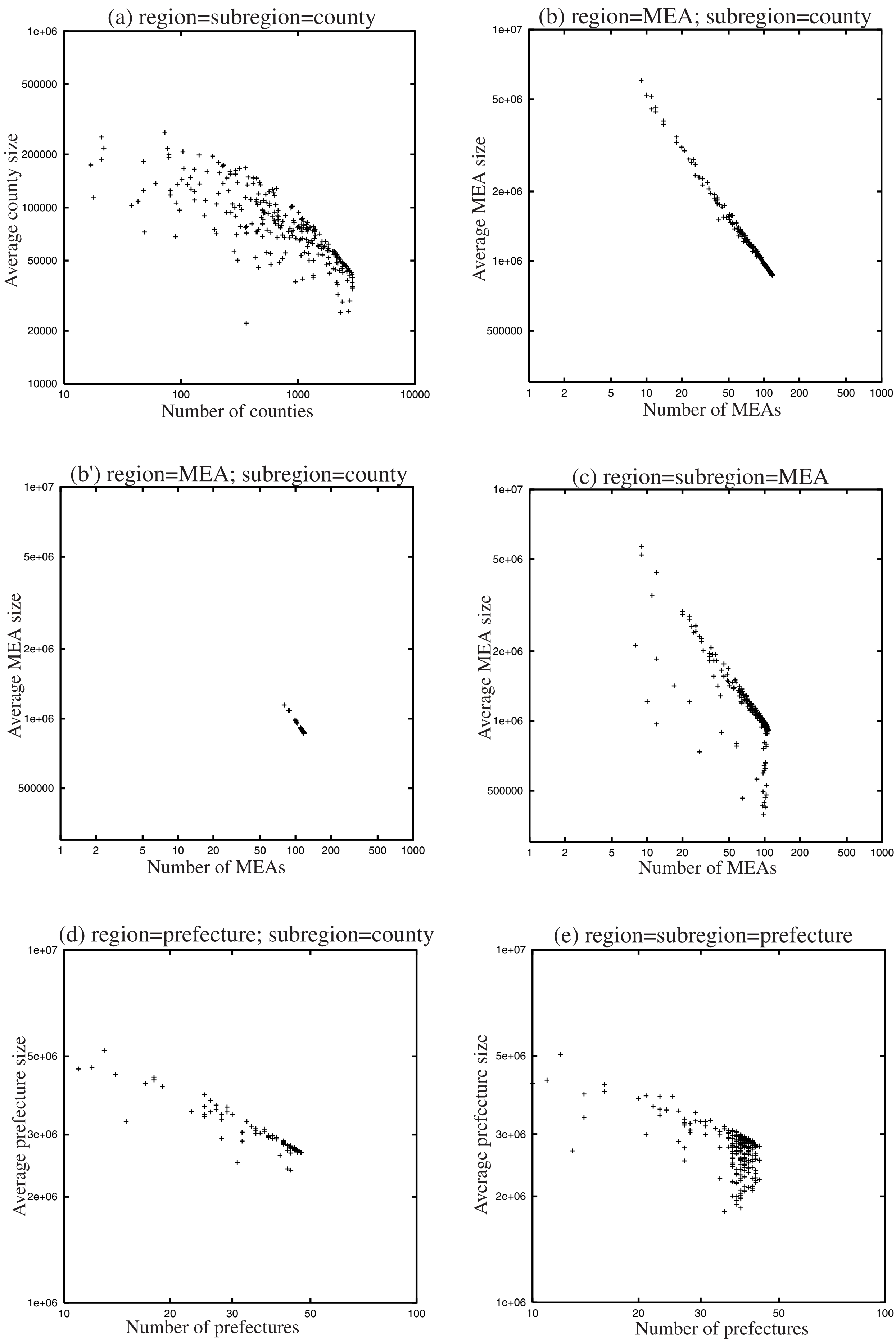


Figure 4. Spatial aggregation levels and the location pattern of industries
(Industry aggregation level is two digit for (b'), and three digit otherwise.)

(a) j is not a neighbor of m

(b) j is a neihbor of m

Figure 5. Neighboring agglomerations

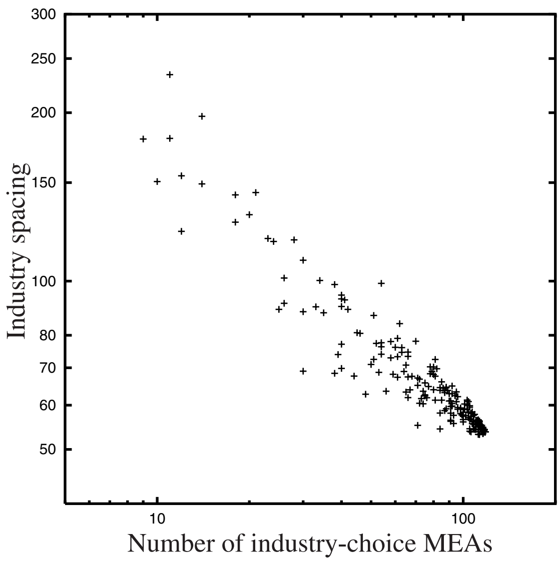


Figure 6. The spacing of industries

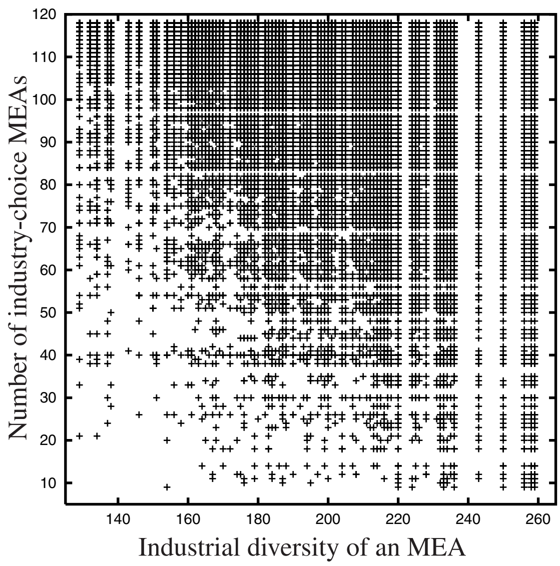


Figure 7. Hierarchical principle

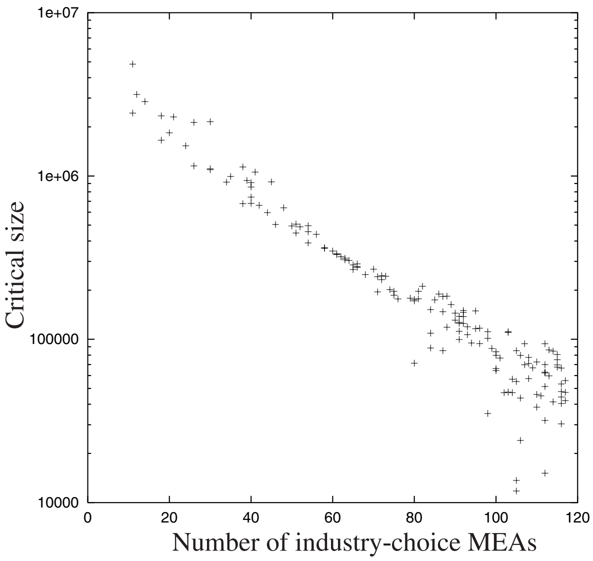


Figure 8. Critical size for industries

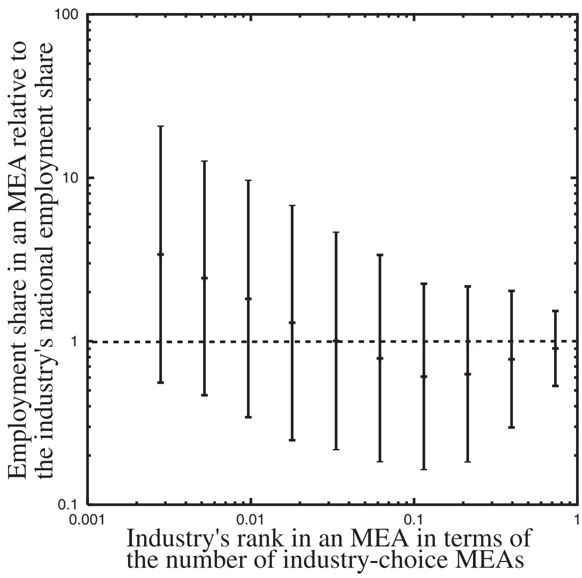


Figure 9. Consistency between hierarchical principle and specialized cities

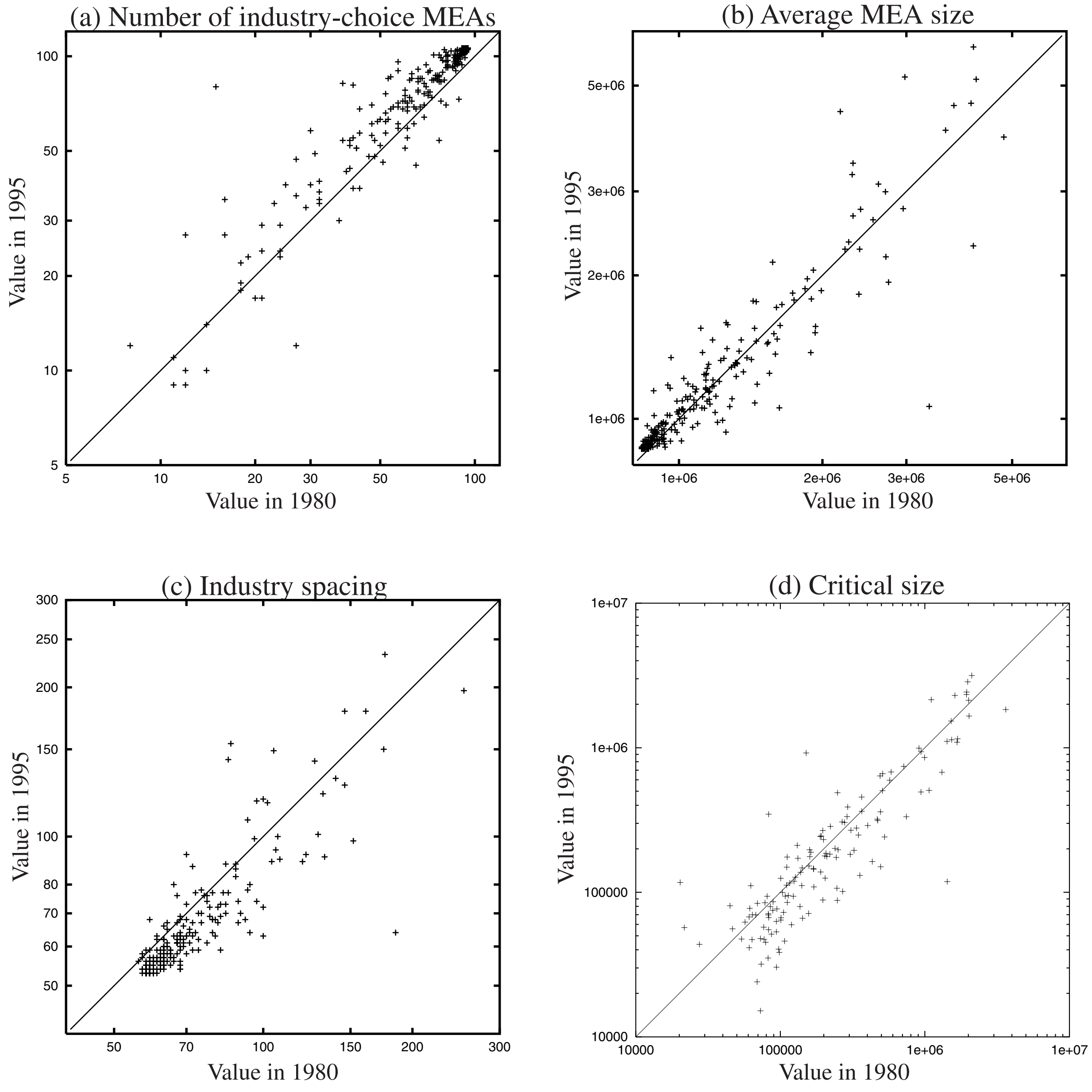


Figure 10. Changes in location pattern of industries.

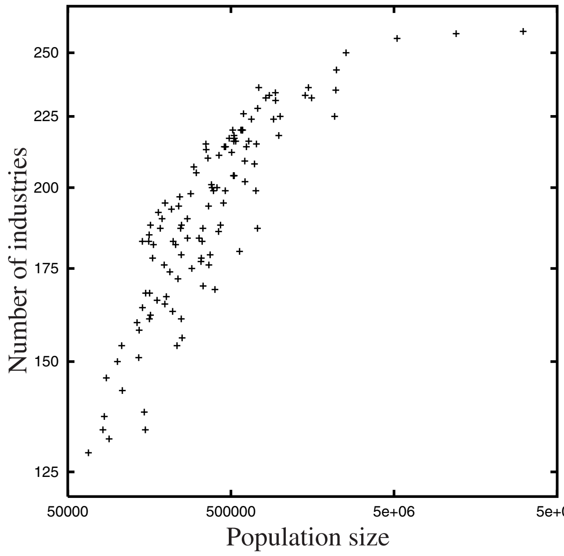


Figure 11. Size and industrial diversity of MEAs

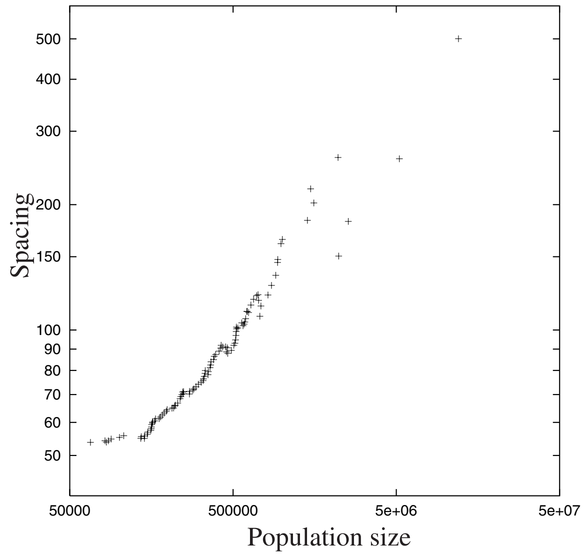


Figure 12. Size and spacing of MEAs

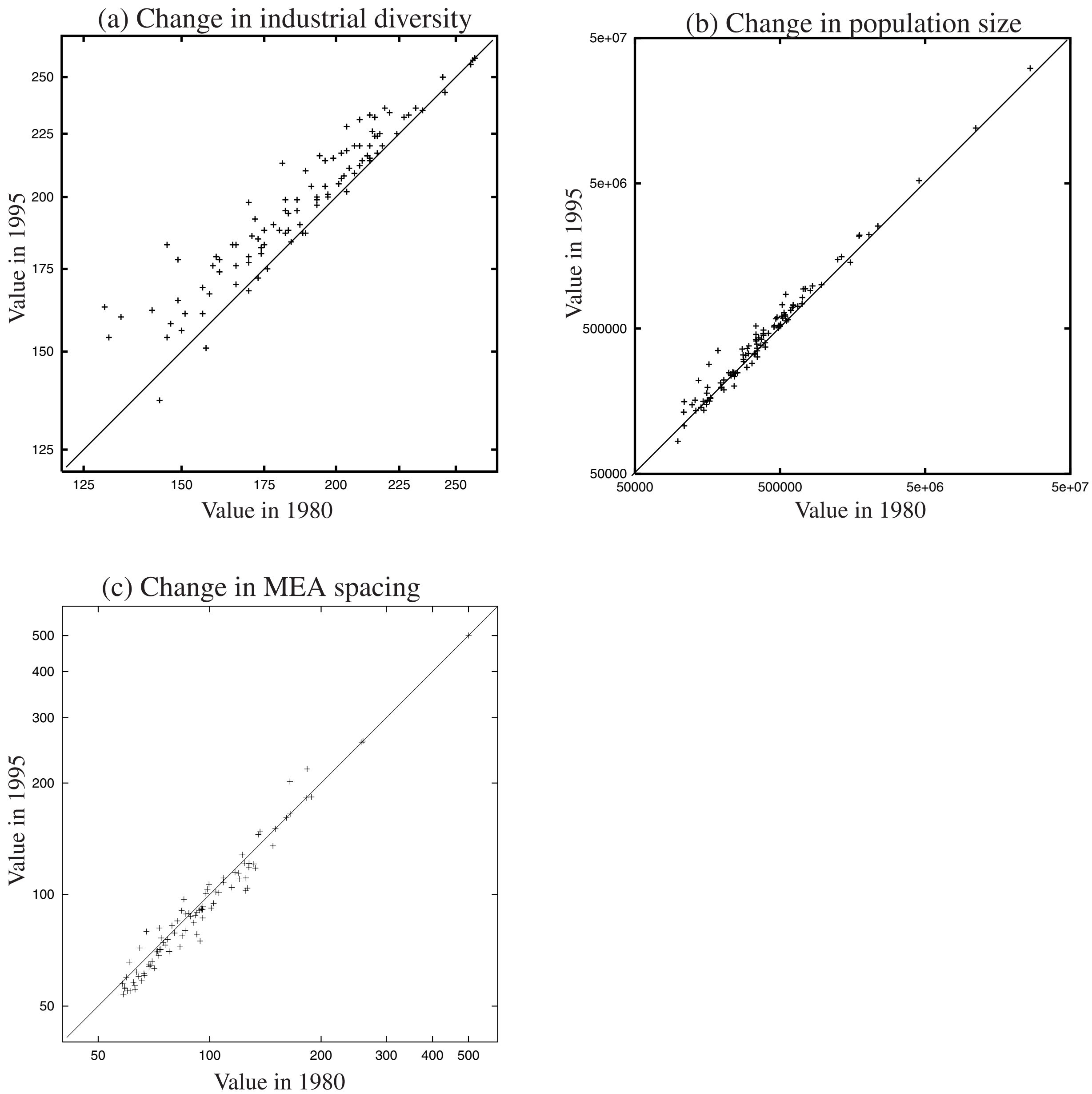


Figure 13. Inter-temporarl change in the industrial diversity, size and spacing of MEAs

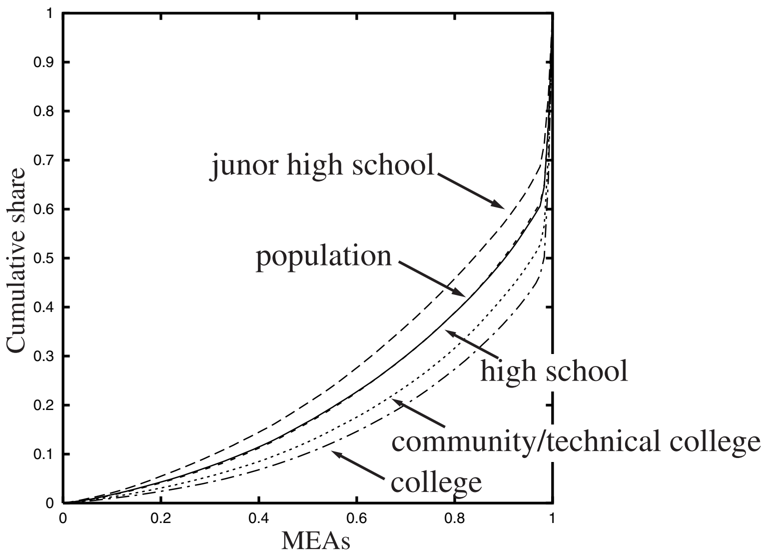
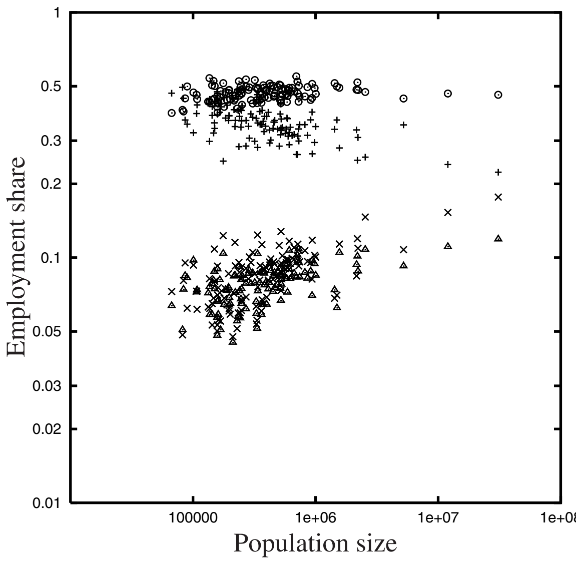
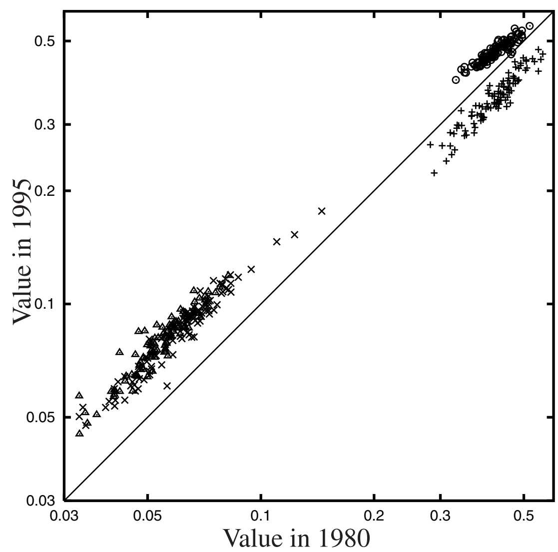


Figure 14. Lorenz curves for spatial distribution of workers with each education level

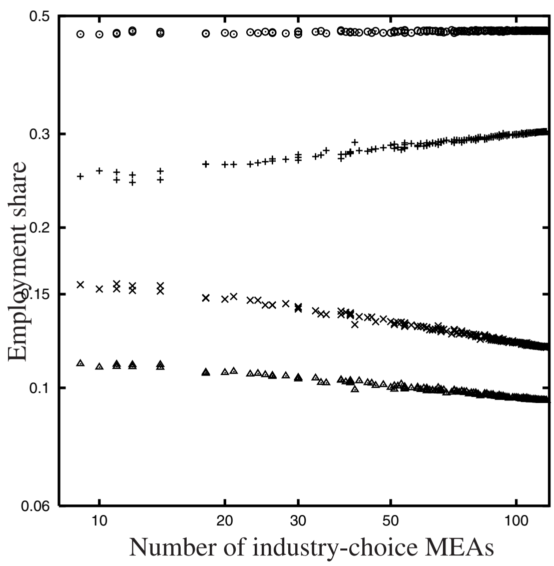


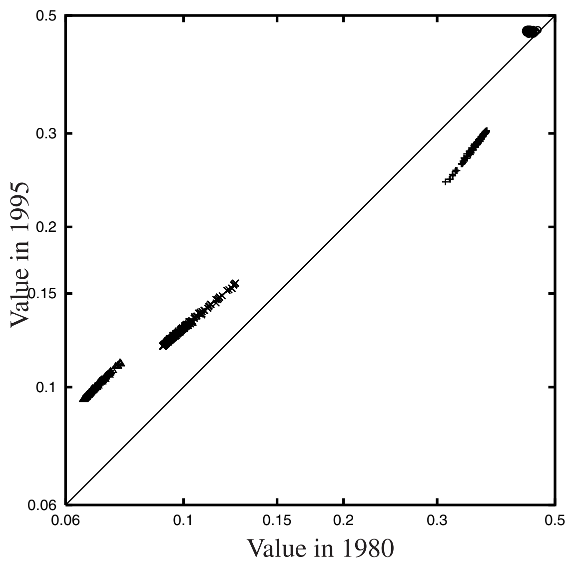


+ Junior high school ∘ High school ▵ Community college × College

Figure 15. Shares of workers in each education level in MEAs

Figure 16. Inter-temporal change in the worker composition of MEAs





+ Junior high school ∘ High school ▵ Community college × College

Figure 17. Worker composition of industries

Figure 18. Inter-temporal change in the worker composition of industries