Spatial patterns of manufacturing clusters in Vietnam

	Gokan Toshitaka, Kuroiwa Ikuo, Nakajima Kentaro, Sakata Shozo
	Copyrights ⓒ / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) http://www.ide.go.jp
journal or publication title	IDE Discussion Paper
volume	616
year	2016-08
URL	http://hdl.handle.net/2344/1583

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 616

Spatial Patterns of Manufacturing Clusters in Vietnam

Toshitaka GOKAN[1], Ikuo KUROIWA[2],

Kentaro NAKAJIMA[3] and Shozo SAKATA[4]
August 2016

Abstract

The formation of industrial clusters is critical for sustained economic growth. We identify the manufacturing clusters in Vietnam, using the Mori and Smith (2013) method, which indicates the spatial pattern of industrial agglomerations using the global extent (GE) and local density (LD) indices. Spatial pattern identification is extremely helpful because industrial clusters are often spread over a wide geographical area and the GE and LD indices—along with cluster mapping—display how the respective clusters fit into specific spatial patterns.

1 Institute of Developing Economies (IDE-JETRO): Toshitaka_Gokan@ide.go.jp

2 Institute of Developing Economies (IDE-JETRO): Ikuo_Kuroiwa@ide.go.jp

3 Tohoku University: knakajima@econ.tohoku.ac.jp

4 Institute of Developing Economies (IDE-JETRO): Shozo_Sakata@ide.go.jp

1. Introduction

The formation of industrial clusters is necessary. In particular, industrial deepening—local linkage formation by creating a robust supplier base and expanding ancillary services (Asian Development Bank 2013)—is necessary for sustained industrial development. The benefits of industrial clusters are multi-faceted. First, the clustering of economic activities reduces the costs of obtaining intermediate inputs and of shipping goods to downstream customers. Second, such clustering induces the emergence of a large pool of labor, which facilitates better matching of workers to employers, and third, information and technology spillovers are more likely (Marshall 1920, Fujita et al. 1999). As a result, firms have a greater incentive to locate near each other, so that industrial agglomerations are shaped to strengthen the competitiveness of industries and sustain industrial development.

Since Vietnam started to liberalize trade and investment, it has joined the global value chains of labor-intensive industries such as apparel and footwear: this was spurred particularly after the signing of the Bilateral Trade Agreement with the USA (BTAUS) in 2000. In recent years, Vietnam's industrial and trade structures have changed significantly. High-tech products, notably mobile phones, have become a leading export product of Vietnam. However, due to the lack of a local supplier base, a large portion of inputs needs to be imported from China and other East Asian countries. Similarly, the automotive industry, which is a typical import-substitution industry in Vietnam, is not internationally competitive because the supplier industry is very weak and only bulky or labor-intensive parts are procured locally. It is obvious that development of the supplier industry and the formation of industrial clusters will be crucially important to enhance the competitiveness of Vietnamese industries.

Against this backdrop, this study identifies the manufacturing clusters in Vietnam, using the method proposed by Mori and Smith (2013). The identified clusters are tested against the hypothesis of spurious clusters. Next, the global extent (GE) and local density (LD) indices—along with cluster mapping—are used to indicate the spatial patterns of industrial clusters. The GE and LD indices are designed to quantify pattern differences in terms of both global and local properties.

Gokan et al. (2015) have already applied the Mori and Smith method to Cambodia,

Laos, and Thailand. Their study shows that the spatial patterns of industrial clusters vary significantly, reflecting the stage of industrial development as well as the geographical properties of the respective countries. In this context, it is important to include Vietnam—an important, rapidly growing economy in the Mekong sub-region—in our study and to further extend the scope of analysis.

The remainder of this paper is organized as follows. Section 2 explains the theoretical background of the method applied in this study. Section 3 presents the results on applying the method of Mori and Smith (2013) to data on Vietnam. Finally, Section 4 concludes with a summary of the findings.

2. Methodology

This section introduces the role and meaning of the two indices introduced by Mori and Smith (2013), as well as the methods used to derive the two indices and to calculate the values of these indices in this study.

2.1. The GE and LD Indices

The index developed by Mori and Smith (2013) was intended to identify the *spatial structure* of industrial agglomeration. For example, the spatial structure in which firms agglomerate at the global (i.e., large spatial) scale but disperse at the local (i.e., small spatial) scale may be observed, whereas the spatial structure in which firms disperse at the global scale but agglomerate at the local scale may also be seen. The use of any single scalar index such as the Gini index, Theil index, and Ellison and Glaeser index (Ellison and Glaeser, 1997) faces the limitation of discerning such differences of spatial structure. Thus, two indices—GE and LD—were introduced by Mori and Smith (2013).

The process for deriving the GE and LE indices can be divided into three steps: first the clusters are detected; second the essential containment is found; and third the GE and LD indices are calculated.

The first step is to find the cluster scheme $\mathbb{C}^*$ that has a maximum value for the Bayes information criteria (BIC) among each candidate cluster scheme:

$$BIC_{\mathbb{C}} = L_{\mathbb{C}}(\hat{P}_{\mathbb{C}}|x) - \frac{k_{\mathbb{C}}}{2}\ln n$$

where

3

$$L_{\mathbb{C}}(\hat{P}_{\mathbb{C}}|x) = \sum_{j=0}^{k_{\mathbb{C}}} n_j(x)\, ln\hat{p}_{\mathbb{C}}(j) + \sum_{j=0}^{k_{\mathbb{C}}} \sum_{r \in C_j} n_r\, ln\frac{a_r}{a_{C_j}}.$$

Cluster scheme $\mathbb{C}$ is simply one or more disjoint clusters (i.e., convex solids, which will be discussed later), C_j, j= 1, ..., k, and the residual set of all non-cluster regions. BIC increases with the larger log-likelihood of $\hat{P}_{\mathbb{C}}$, that is, the location probability of cluster scheme $\mathbb{C}$ for each basic region, given an observed location pattern x; conversely, BIC decreases with the penalty term composed by $k_{\mathbb{C}}$, which denotes the number of clusters in a cluster scheme $\mathbb{C}$, and n, the number of establishments in the total area.

The location probability of cluster scheme $\mathbb{C}$ for cluster j= 1, ..., k, $p_{\mathbb{C}}(j)$ can be rewritten as $\hat{p}_{\mathbb{C}}(j) = n_j(x)/n$, where $n_j(x)$ denotes the sum of the number of establishments in cluster j= 1, ..., k, and n is the total number of establishments in the whole region. In the above log-likelihood functions, $n_j(x)$ and n_r are related with a sequence of independent location decisions by individual establishments. Because each region is included as part of a certain cluster C_j, j= 1, ..., k, or the residual set of all non-cluster regions, the right hand side of the log-likelihood function of the location probabilities (i.e., the probability that a randomly sampled establishment locates in a region within a certain cluster) expresses the law of total probability: the log-likelihood function can be divided into two parts, namely the first term, which gives the location probabilities that n_j establishments locate in a cluster C_j, j= 1, ..., k, and the second term, which gives the location probability that n_r establishments locate in region r in cluster C_j, j= 1, ..., k, given that individual establishments choose their location completely randomly within each cluster. The location probability of region r under the condition that an establishment locates in a cluster C_j, j= 1, ..., k, is a_r/a_{C_j}, where a_r denotes the economic area in region r and a_{C_j} represents the economic area in a cluster C_j, j= 1, ..., k. The derivation of the economic area is explained in section 2.3. In the process of choosing an additional region within a cluster, the length of the shortest path is used.

The second step is divided into two parts. The first part is to find the essential clusters, which are the most significant clusters in terms of incremental contributions to

BIC under the condition that the sum of incremental contributions to BIC exceeds a certain proportion (λ) of $BIC_\mathbb{C}$. The last part of the second step is to find the smallest convex-solid set in the total area containing the set of essential clusters. Intuitively, a convex-solid set means that the set is connected and does not only have no dents in its perimeter, but also does not have any internal cavities. Then, the regions in the smallest convex-solid set can be regarded as an essential containment.

Finally, we can obtain the GE for an industry by dividing the total economic area of essential containment for an industry by the total economic area of the whole country. Likewise, we can obtain the LD for the industry by dividing the total economic area of the essential clusters for the industry by the total economic area of the essential containment.

2.2. A Test of Spurious Clusters

The recent trend in the index of industrial agglomeration is to implement the test against the null hypothesis that spatial distribution could simply have emerged by chance (Ellison and Glaeser, 1997; Duranton and Overman, 2005; and Mori, *et al.*, 2005). In Mori and Smith (2013), a test of spurious clusters was conducted to test whether the BIC of the best cluster scheme is significantly better than that under a random location pattern generated by a Monte Carlo test.

2.3. The Derivation of the Global Extent and Local Density on Vietnam

To obtain the GE and LD indices for Vietnam, we have applied the above-mentioned method to the data with respect to distance, economic area, and the number of establishments.[5] Basic regions, which we mentioned above, are taken to be the second administrative level, that is, district-level subdivisions of Vietnam. Considering the consistency within the regions in the census data, and after removing islands, we focus on 649 regions. Economic regions are calculated by subtracting the areas of forests, lakes, marshes, and undeveloped areas from the total area of the region, using a satellite

---

[5] The computer programs for these analyses introduced by Mori and Smith (2013) are available at http://www.mori.kier.kyoto-u.ac.jp/data/cluster_detection.html.

image in Global Land Cover 2000. The travel distance between each pair of neighboring regions is estimated by calculating the shortest distance between the centroid of each region on a rod network—for which the information is obtained from a shape file. The shortest-path distances are computed as in Mori and Smith (2014). After omitting any area that is not suitable for economic activity, we calculate the economic area of each region using the shape files. The explanation of the establishment data is outlined in the following sections. Finally, following the method recommended by Mori and Smith (2014), we have chosen the value of λ appropriately at 0.91.

2.4. The Combination of the GE and LD Indices

The combination of the GE and LD indices reflects the *spatial structure* of industrial agglomeration; this follows because the value of GE indicates the share of the total area of essential containment (black and grey squares in Figure 1-1)—which is the smallest convex-solid set containing the set of essential clusters—relative to that of the entire country (the entire area in Figure 1-1). The value of LD expresses the share of the total area of its essential clusters (black squares in Figure 1-2) relative to that of its essential containment (black and grey squares in Figure 1-1 or 1-2).

- Figure 1 -

Thus, a larger GE may represent the spatial distribution of an industry that is more dispersed throughout a country, whereas a larger LD may reveal the spatial distribution of an industry that is more dispersed throughout the essential containment. In other words, industries with a low GE are regarded as exhibiting a relatively "confined" spatial pattern, while those with a high GE demonstrate a relatively "dispersed" pattern. Likewise, industries with a high LD indicate a relatively "dense" pattern, while those with a low LD demonstrate a relatively "sparse" pattern. Thus, we obtain four patterns of spatial structure, namely the "globally dispersed and locally dense" pattern; the "globally dispersed and locally sparse" pattern; the "globally confined and locally dense" pattern; and the "globally confined and locally sparse" pattern. These patterns are used in Section 3.

6

To get a clearer image of the four patterns of spatial structure, we exemplify the GE and LD indices using a box composed of 36 blocks. We present the resulting four simplified figures for the GE and LD indices for each pattern in Figure 2; λ is set to be 1 for simplicity. Each of the four boxes has 36 blocks, which implies that there are 36 regions in the entire country. The number of black and grey regions is used to calculate the GE and LD indices. The black regions represent the regions in an essential cluster. The grey regions are not essential clusters, but are included in the essential containment.

Figure 2-1 indicates the globally confined and locally dense pattern. From the number of regions of the total area, the essential containment and essential cluster are 36, 9, and 8, respectively; we calculate the GE as 9/36=0.25 and the LD as 8/9=0.89. In the same manner, we can obtain the GE and LD indices for the other three spatial structure patterns.

Notice the differences in the spatial structure, although the numbers of the regions contained in the essential clusters are identical in Figure 2-1 and Figure 2-3. The usage of the GE and LD indices clarifies the differences in spatial structure.

- Figure 2 -

3. Empirical Results

This section introduces the empirical results of the analyses of spatial patterns of industrial clusters that are introduced in Section 2.

We used the data of the fourth (2012) Establishment Census conducted by the General Statistics Office (GSO). The census covers all economic enterprises including state, non-state, and foreign-invested sectors; it was conducted on 1 April 2012, collecting information as of 31 December 2011. The country has 341.6 thousand enterprises as of 31 December 2011. The data that we used for this study include the location of the establishment (district level), establishment code (ISIC 4 digit), and number of laborers.

3.1. Four Location Patterns of Industrial Clusters

The GE and LD indices—and the combination of these two indices—indicate the spatial

patterns of industrial clusters. As shown in Figure 3, the four location patterns are demonstrated in line with Mori and Smith (2013).

- Figure 3 -

(1) Globally Dispersed and Locally Dense Patterns

Industries with relatively high values for GE and LD exhibit globally dispersed and locally dense patterns of industrial agglomeration: these industries are located in the upper right corner of Figure 3. One such example is the manufacture of starches and starch products (1062: GE = 0.46 and LD= 0.61).[6] This sector includes the manufacture of rice noodles, which is a national dish in Vietnam. Therefore, it is quite natural that the manufacturers of this sector are spread all over the country. As a result, both the industrial clusters and the essential containment cover a broad area (see Figures 4-2 and 4-3).

- Figure 4 -

 (2) Globally Dispersed and Locally Sparse Patterns

Industries with relatively high values of GE and low values of LD—which are positioned near the southeast corner in Figure 3—exhibit globally dispersed and locally sparse patterns. The manufacture of malt liquors and malt (1103: GE=0.66, LD=0.14) is such an example. As shown in Figure 5-1, malt liquor is produced in a broad area that includes the northern region; in addition, the essential containment covers a large area (Figure 5-3). However, unlike starches and starch products, malt liquor clusters are concentrated in a smaller number of districts (Figure 5-2).

- Figure 5 -

---

[6] 1062 and other code numbers are in accordance with the Vietnam Standard Industrial Classification (VSIC).

(3) Globally Confined and Locally Dense Patterns

The manufacture of grain mill products (1061: GE=0.06, LD=0.83) is an example of globally confined and locally dense patterns. As shown in Figure 6-2, clusters of grain mill products are concentrated in a relatively small area of Southern Vietnam, while the essential containment is concentrated in Ho Chi Minh and the Mekong Delta region (Figure 6-3).[7]

- Figure 6 –

(4) Globally Confined and Locally Sparse Patterns

Industries with relatively low values of GE and LD—which are positioned near the southwest corner in Figure 1—exhibit globally confined and locally sparse patterns. An example of this location pattern is provided by the manufacture of railway locomotives and rolling stock (3020: GE=0.12, LD=0.007).

- Figure 7 -

- Figure 8 -

It should also be noted that many clusters are concentrated in a narrow range where its GE is around 0.2 and its LD is less than 0.2; more specifically, 75 sectors (out of 137 sectors) are in ranges of 0.11–0.22 (GE) and 0.005–0.19 (LD), respectively. Manufacturing industries such as the subdivision or sub-subdivision of spinning, weaving, and finishing of textiles; wearing apparel; leather and related products; pharmaceuticals, medicinal chemical and botanical products; rubber and plastics products; electrical equipment; general purpose machinery; and other manufacturing fall into this category. In addition, the food and beverage industries (such as bakery

---

[7] Industries with low values of GE and high values of LD—which are positioned near the northwest corner in Figure 1—exhibit globally confined and locally dense patterns. The manufacture of wearing apparel, except fur apparel (NEC: not elsewhere classified), appears to be the most suitable example because its GE value is near zero and its LD value is one. However, the manufacture of wearing apparel, except fur apparel (NEC), contains only one establishment that automatically produces the abovementioned values of GE and LD. It is obvious that such an extreme case—where there is only one establishment in a sector—does not meet the criteria and can therefore not be considered as an industrial cluster.

products; prepared meals and dishes; and wines and tobacco products) are included in this category. Mapping of these clusters shows that they are mainly located in and around Hanoi and/or Ho Chi Minh City.

4. Conclusion

We applied the Mori and Smith (2013) method to identify manufacturing clusters in 137 manufacturing sectors of Vietnam. The establishment-level official statistics enabled us to categorize the spatial distribution of establishments into four spatial structures.

In Vietnam, the starches and starch products sector exhibits a globally dispersed and locally dense spatial pattern. Since this sector includes the manufacture of rice noodles, which is a national dish in Vietnam, it reflects the fact that staple foods (and their materials) are cultivated and consumed ubiquitously within a country, so that they tend to exhibit a globally dispersed and locally dense spatial pattern.[8]

The malt liquors and malt sector exhibits a globally dispersed and locally sparse spatial pattern. Considering the characteristics of these industries, such a location pattern can be observed when the consumers of these products are spread across a broad area; however, the production process requires factors that are not available ubiquitously: such factors include raw materials and workers' skills that are only available locally.[9]

The grain mill products sector exhibits a globally confined and locally dense spatial pattern. This spatial pattern is observed because grain mill clusters are concentrated in a relatively small area of Southern Vietnam: in particular, the essential containment is concentrated in Ho Chi Minh and the Mekong Delta region.[10]

A globally confined and locally sparse spatial pattern indicates that there are a

---

[8] Similarly, in all of Cambodia, Lao PDR, and Thailand, the grain and rice-milling sector exhibits a globally dispersed and locally dense spatial pattern (Gokan et al. 2014).
[9] A globally dispersed and locally sparse spatial pattern is observed in dairy products (Cambodia), frozen meat of mammals (Lao PDR), as well as fish products, wine, glass products, ceramic ware, and so on (Thailand) (Gokan et al. 2014).
[10] There are many industries that exhibit a globally confined and locally dense pattern in Cambodia and Lao PDR. Industries that contain only one establishment automatically illustrate this location pattern. The underdevelopment of manufacturing industries is reflected in such a location pattern of these economies, whereas the case of Thailand attributes economies of scale in production as one of the reasons for exhibiting this location pattern (Gokan et al. 2014).

limited number of industrial clusters that constitute a relatively narrow essential containment. In the case of Vietnam, the manufacture of railway locomotives and rolling stock forms such a cluster.[11] It is worth to note that many clusters are concentrated in a narrow range where its GE is around 0.2 and its LD is less than 0.2. These clusters are mainly located in and around Hanoi and/or Ho Chi Minh City.

Finally, it is important to point out several limitations of our study. The first is the lack of analysis on the collocation pattern of different industries. Agglomeration of an industry can attract other industries through agglomeration externalities. This limitation makes it difficult for a policymaker to predict which clusters will induce the development of which new industries. The second is the usage of cross-sectional data. Panel data are needed to observe whether agglomeration is associated with industrial growth by the existing industries, or with the generation of new industries.

---

[11] Various industries exhibit a globally confined and locally sparse pattern: in Cambodia, pulp, paper, and paperboard; in Lao PDR, other knitted or crocheted fabrics; and in Thailand, leather; recorded media; steam generators; instruments and appliances for measuring, checking, testing, and other purposes; as well as aircraft and spacecraft (Gokan et al. 2014).

References

Asian Development Bank (2013), "Asia's Economic Transformation: Where to, How, and How Fast?" *Key Indicators for Asia and the Pacific.* Part 1 Special Chapter.

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

Ellison, G. and E. L. Glaeser (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy,* 1997, 105(5), pp. 889–927.

Fujita, M. and J.-F. Thisse (2002) *Economics of Agglomeration Cities, Industrial Location and Regional Growth* Cambridge: Cambridge University Press.

Gokan, T., I. Kuroiwa, N. Laksanapanyakul, and Y. Ueki (2015), "Spatial Patterns of Manufacturing Clusters in Cambodia, Lao People's Democratic Republic, and Thailand." ERIA Discussion Paper 2015-68.

Marshal, A. (1890) *Principle of Economics*, http://www.econlib.org/library/Marshall /marP.html.

Mori, T. and T. E. Smith (2013), "A Probabilistic Modeling Approach to the Detection of Industrial Agglomerations," *Journal of Economic Geography*, 14(3), pp. 547– 588.

Mori, T. and T. E. Smith (2014), "On the Spatial Scale of Industrial Agglomerations," *Journal of Urban Economics*, forthcoming.

Mori, T., K. Nishikimi, and T. E. Smith (2005), "A Divergence Statistic for Industrial Localization," *Review of Economics and Statistics,* 87, pp. 635–651.

Figure 1: Calculation of Global Extent and Local Density

(1-1) Calculation of Global Extent
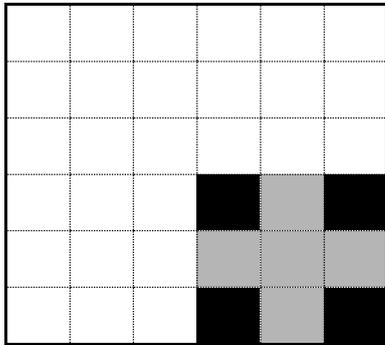
(1-2) Calculation of Local Density

Source: Drawn by the authors
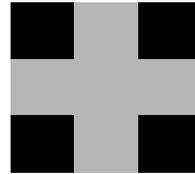
Figure 2: Four Combinations of Global Extent and Local Density

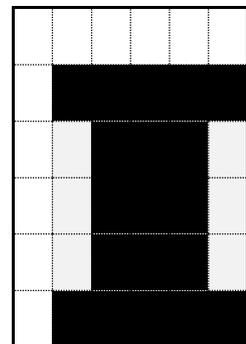(2-1) Globally confined and locally dense pattern
GE = 9/36, LD = 8/9
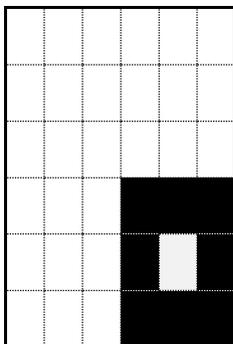
(2-2) Globally confined and locally sparse pattern
GE = 9/36, LD = 4/9
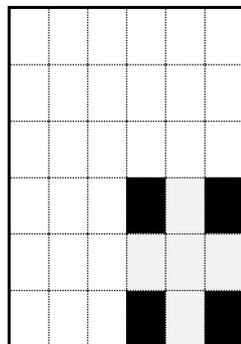
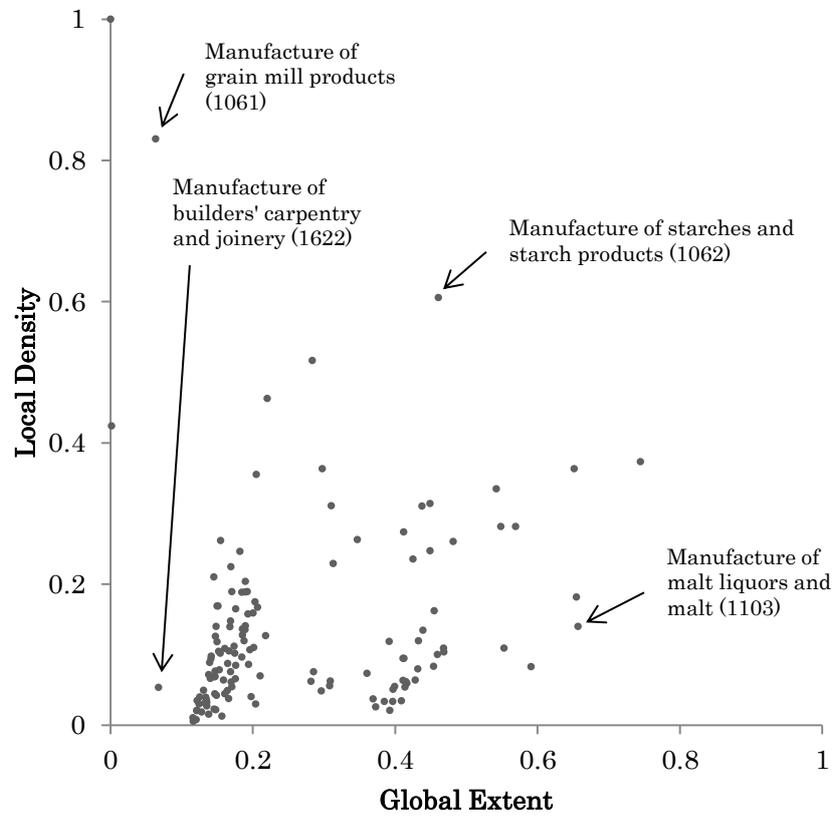(2-3) Globally dispersed and locally sparse pattern
GE = 25/36, LD = 8/25
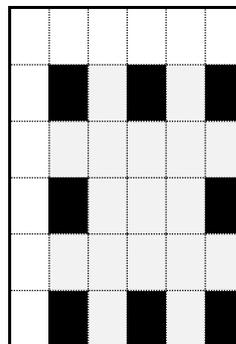
(2-4) Globally dispersed and locally dense pattern
GE = 25/36, LD = 19/25

Note : GE = global extent, LD = local density.

Source: Drawn by the authors

Figure 3: Global Extent and Local Density in Vietnam (2011)

Source: Drawn by the authors

Figure 4: Manufacture of starches and starch products (1062)
(4-1) Density of establishments (per km^2)

Density of establishments (per square km)

	0.000000 - 0.000971
	0.001517 - 0.003964
	0.005000 - 0.007984
	0.012500 - 0.014235
	0.023071 - 0.031332
	0.041885 - 0.051948
	0.070175 - 0.082148
	0.242424 - 0.347826
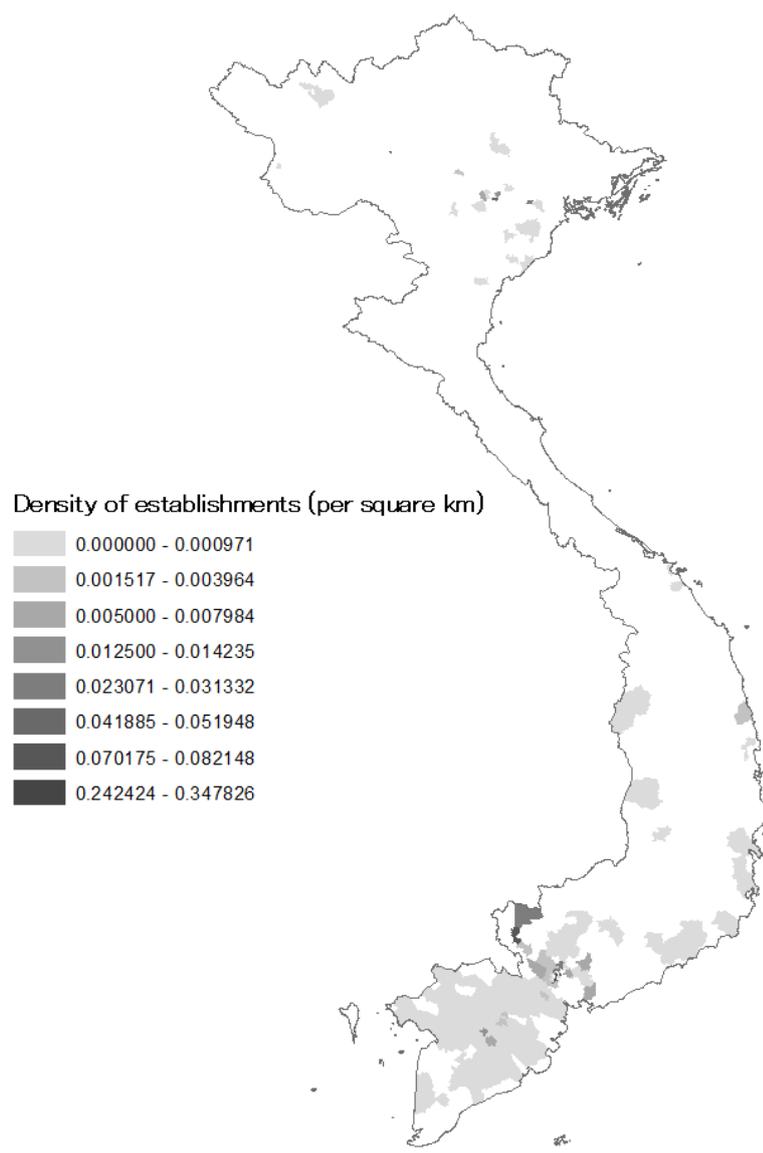
Source: Drawn by the authors

Figure 4: Manufacture of starches and starch products (1062)
(4-2) Location of clusters
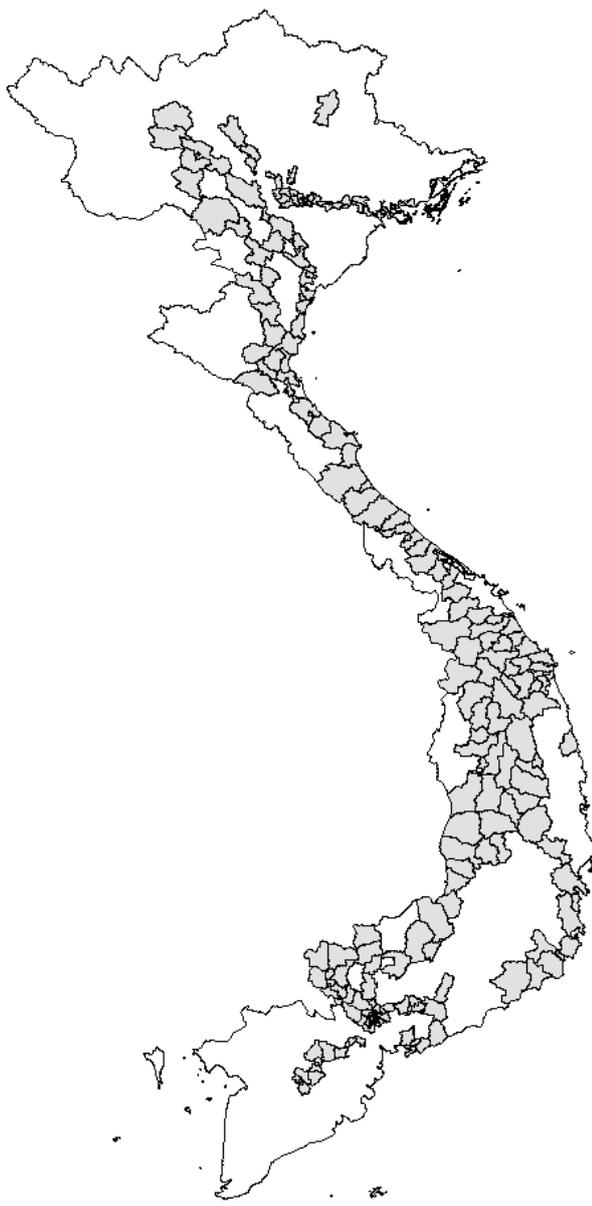
Source: Drawn by the authors

Figure 4: Manufacture of starches and starch products (1062)
(4-3) Essential containment
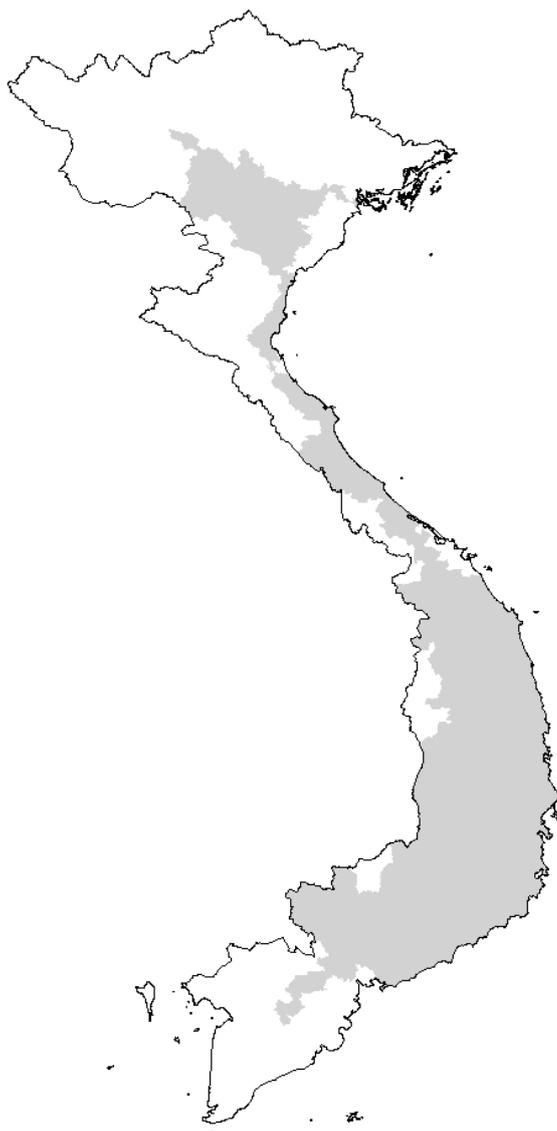
Source: Drawn by the authors

Figure 5: Manufacture of malt liquors and malt (1103)
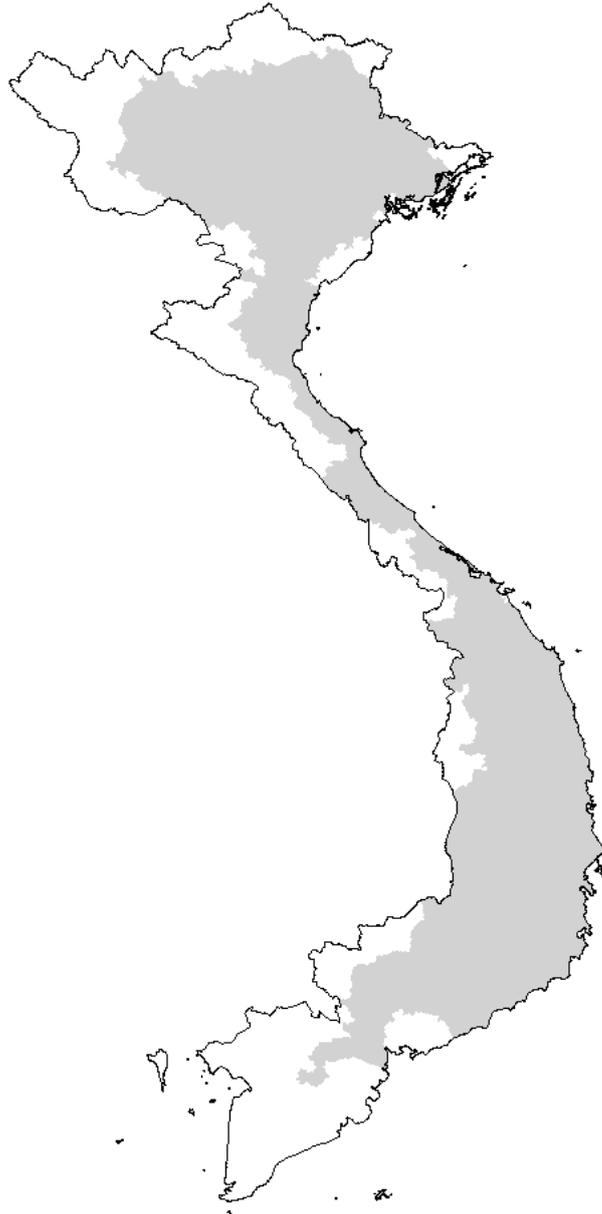**(5-1) Density of establishments (per km^2)**

Density of establishments (per square km)
- 0.000971 - 0.003850
- 0.003851 - 0.007319
- 0.007320 - 0.011461
- 0.011462 - 0.019139
- 0.019140 - 0.035398
- 0.035399 - 0.061069
- 0.061070 - 0.200000
- 0.200001 - 0.400000
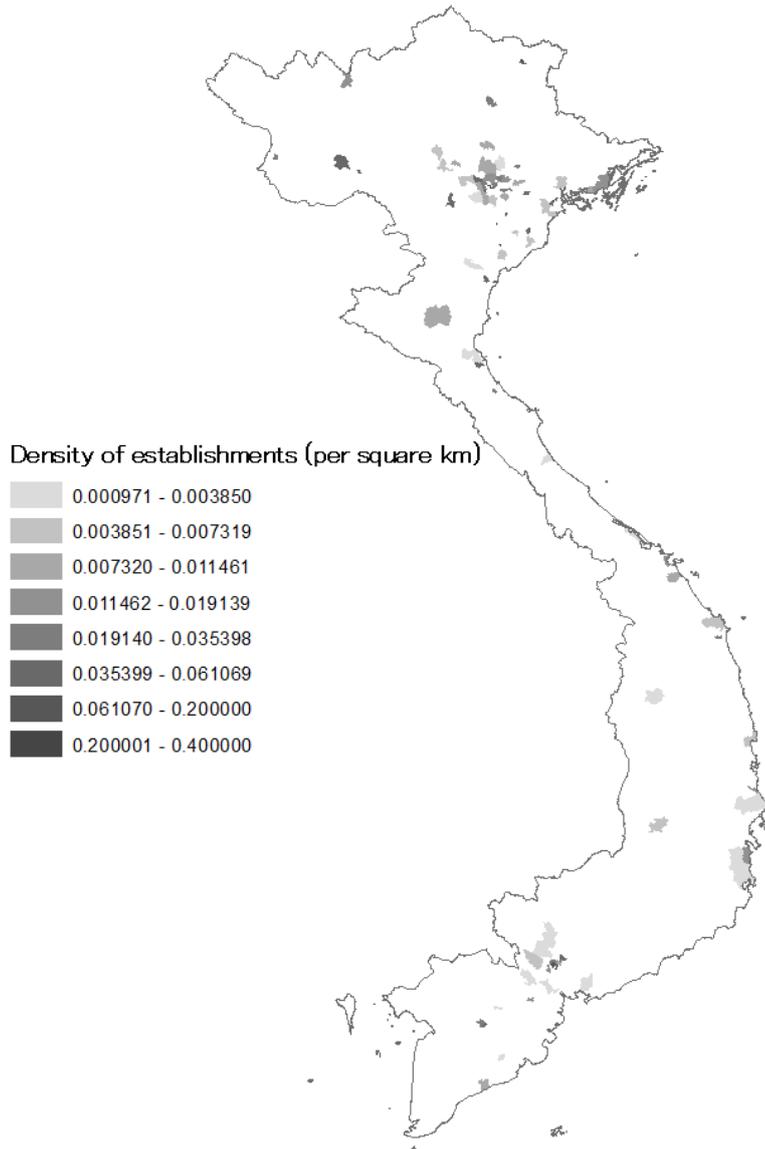
Source: Drawn by the authors

Figure 5: Manufacture of malt liquors and malt (1103)
(5-2) Location of clusters
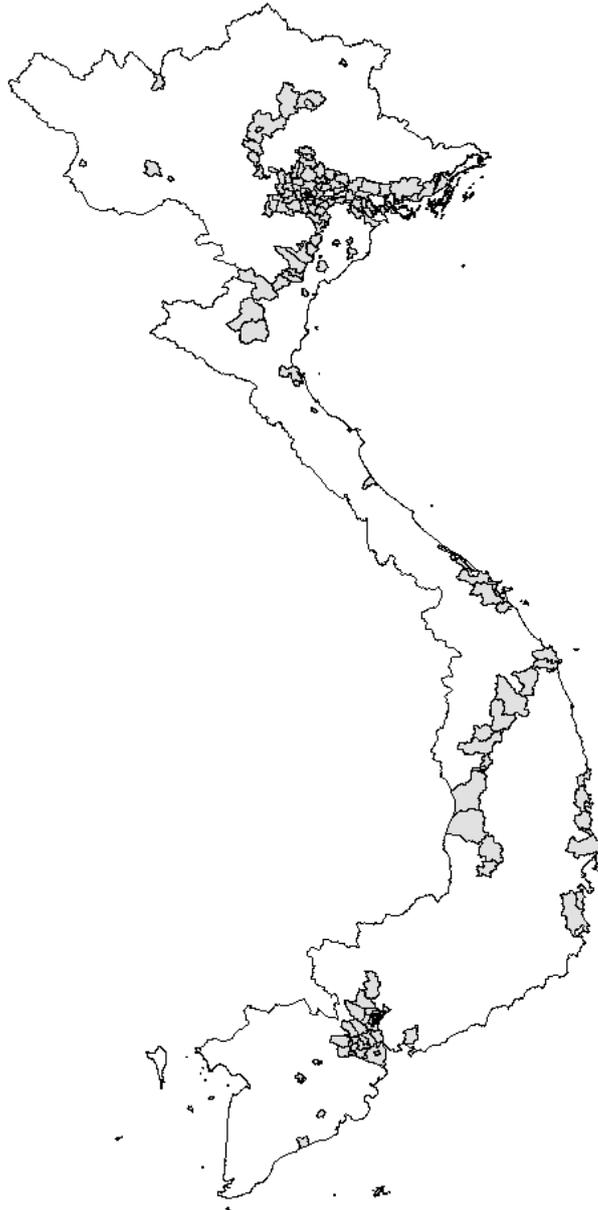
Source: Drawn by the authors

Figure 5: Manufacture of malt liquors and malt (1103)
(5-3) Essential containment

Source: Drawn by the authors

Figure 6: Manufacture of grain mill products (1061)
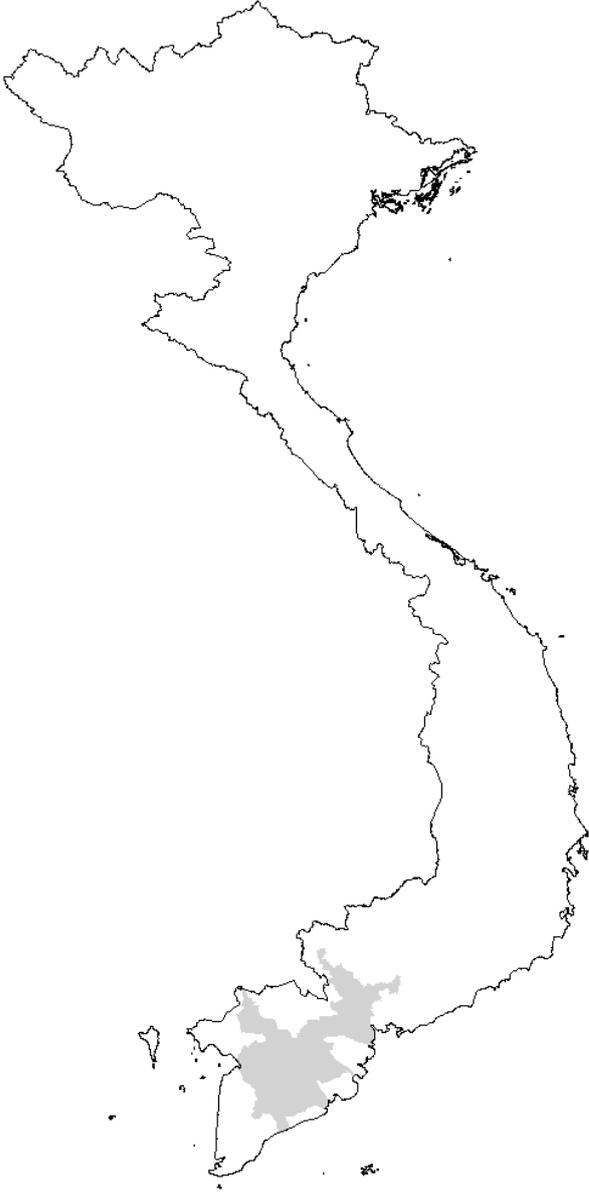**(6-1) Density of establishments (per km^2)**

Density of establishments (per square km)

	0.001337 - 0.005626
	0.005801 - 0.013005
	0.015038 - 0.029304
	0.031688 - 0.053333
	0.057942 - 0.101911
	0.126316 - 0.192077
	0.216216 - 0.296296
	0.363636 - 0.569343
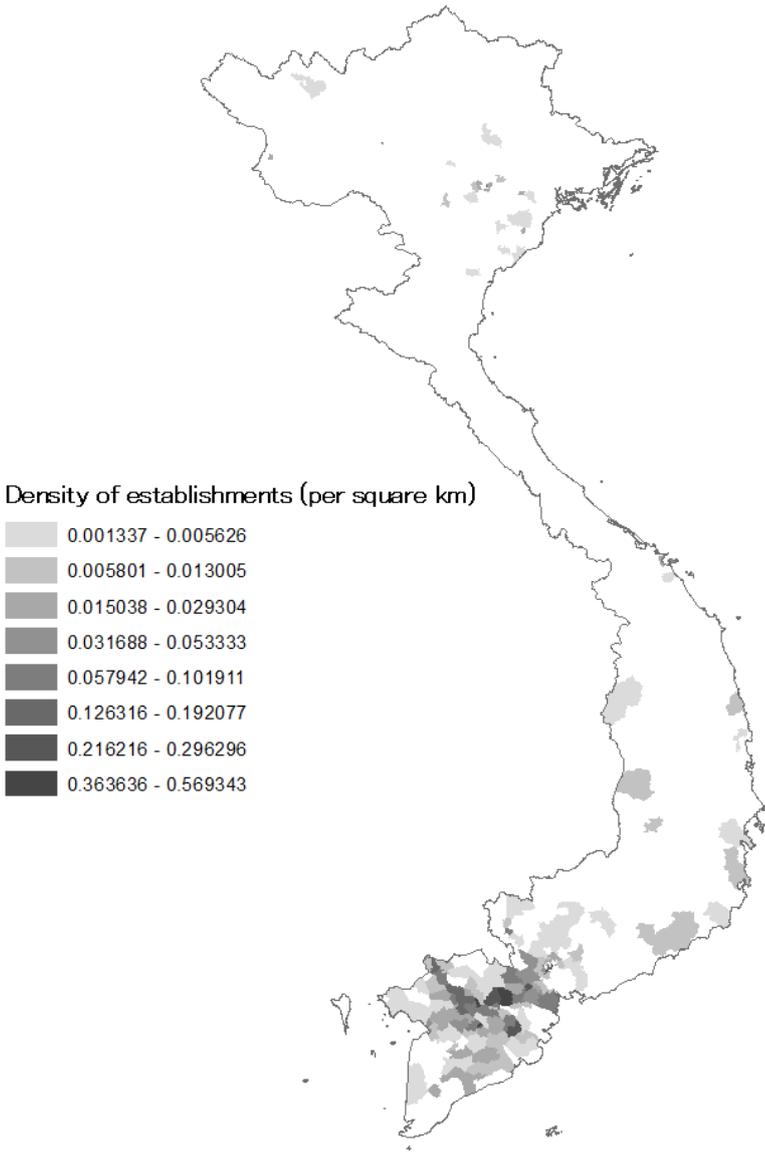
Source: Drawn by the authors

Figure 6: Manufacture of grain mill products (1061)
(6-2) Location of clusters
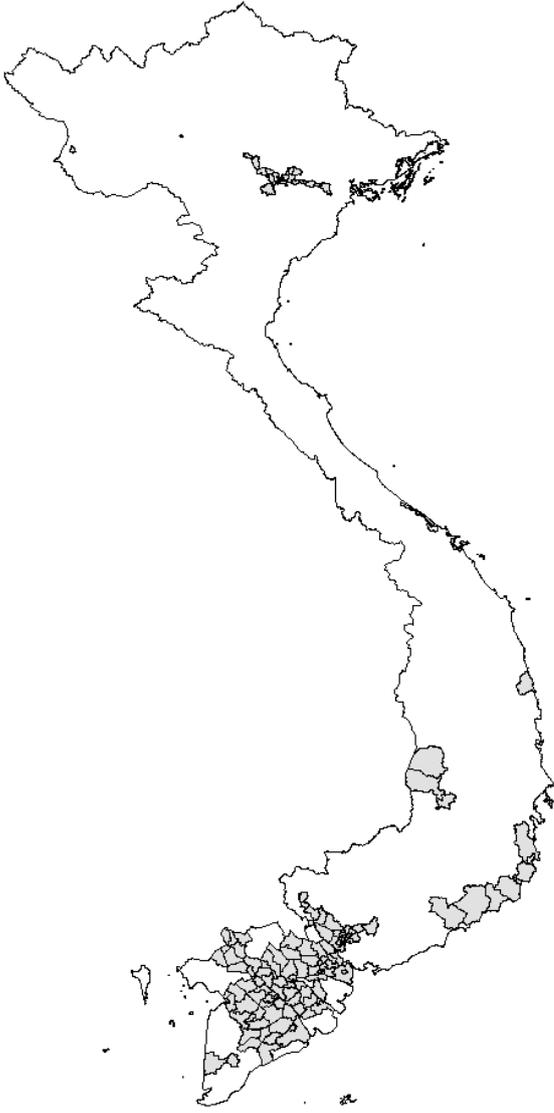
Source: Drawn by the authors

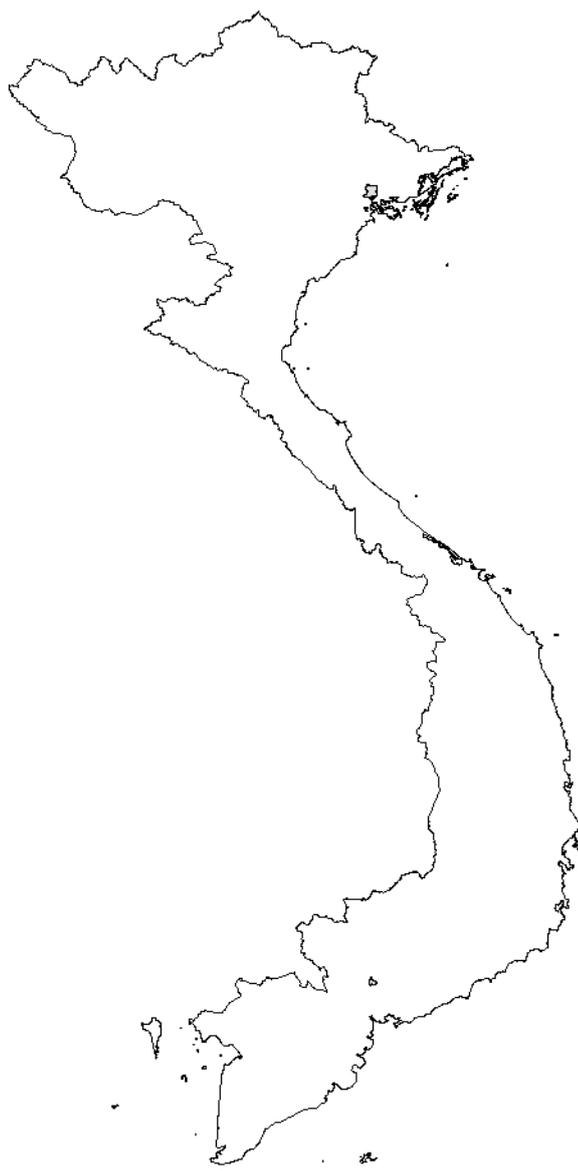**Figure 6: Manufacture of grain mill products (1061)**
(6-3) Essential containment

Source: Drawn by the authors

Figure 7: Manufacture of railway locomotives and rolling stock (3020)
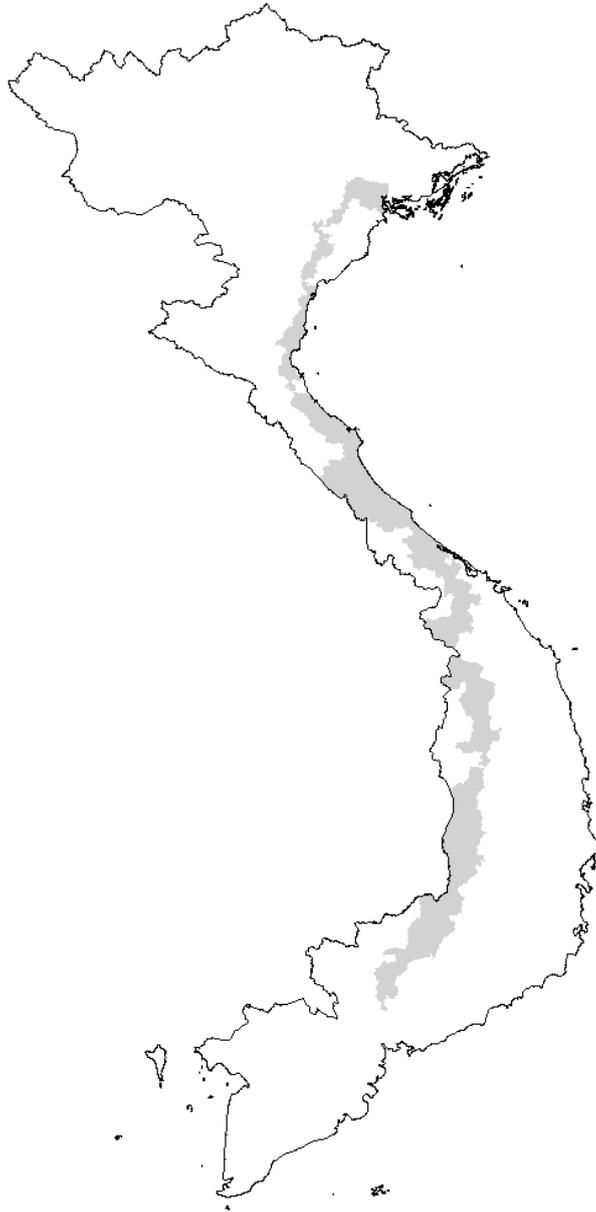**(7-1) Density of establishments (per km^2)**
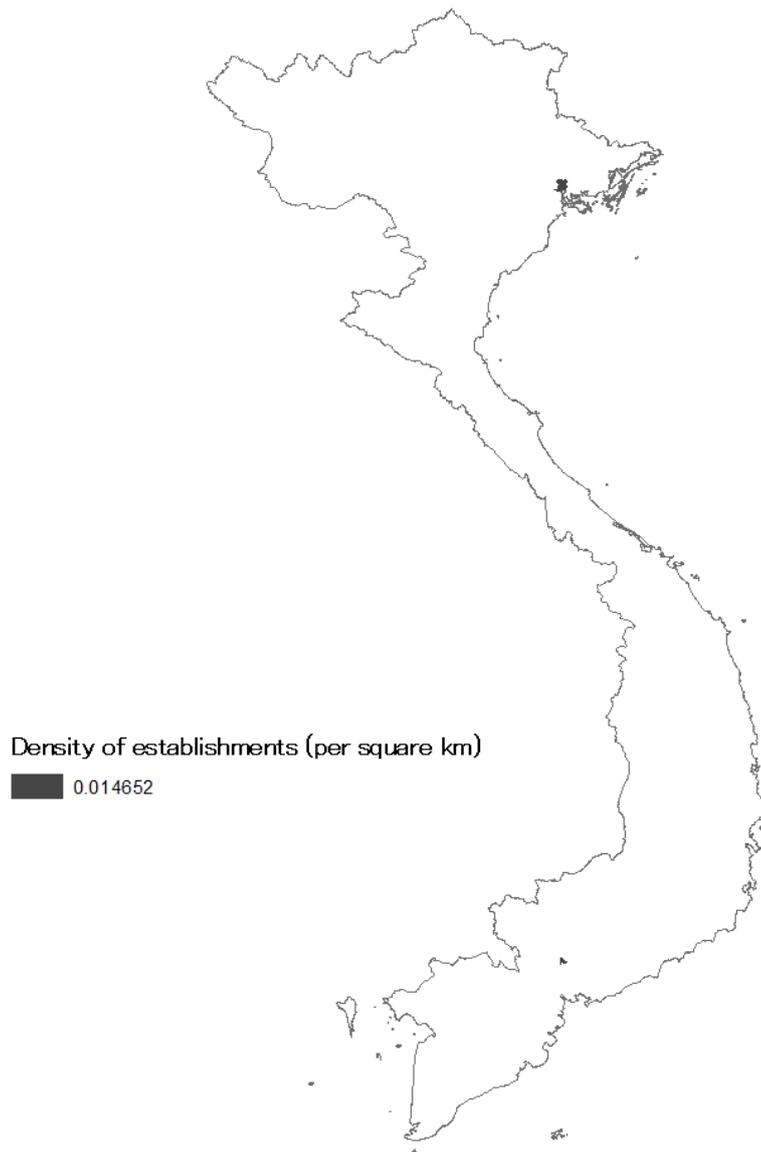
Density of establishments (per square km)

0.014652

Source: Drawn by the authors

Figure 7: Manufacture of railway locomotives and rolling stock (3020)
(7-2) Location of clusters

Source: Drawn by the authors

Figure 7: Manufacture of railway locomotives and rolling stock (3020)

(7-3) Essential containment

Source: Drawn by the authors

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.