RIETI Discussion Paper Series 11-E-076

How Important is Geographical Agglomeration to Factory Efficiency in Japan's Manufacturing Sector?

FUKAO Kyoji
RIETI

Victoria KRAVTSOVA
Hitotsubashi University

NAKAJIMA Kentaro
Tohoku University

RIETI
Research Institute of Economy, Trade & Industry, IAA

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

RIETI Discussion Paper Series 11-E-076
December 2011

How Important is Geographical Agglomeration to Factory Efficiency in Japan's Manufacturing Sector?

FUKAO Kyoji
Research Institute of Economy, Trade and Industry / Hitotsubashi University
Victoria KRAVTSOVA[1]
Hitotsubashi University
NAKAJIMA Kentaro
Tohoku University

Abstract

In this paper, geographical spillover potential is modeled and empirically examined using factory-level data from Japan's Census of Manufactures. First, the efficiency of each factory is estimated using a non-parametric data envelopment analysis (DEA) model for each industry. Second, the geographical distances to the most efficient factory in the prefecture and Japan overall are estimated. Third, the determinants of the factories' performance are identified and estimated. We find that clustering occurs in each industry, and efficient factories concentrate in certain regions. The percentage of efficient firms out of the total number of firms is particularly high in the Chubu and Tohoku regions. The estimation results also suggest that proximity to the most efficient factories plays a statistically significant role in determining the efficiency of factories in Japan in most industries. However, this is not the case in high-tech industries.

Key words: regional policy; spatial models; productivity and efficiency.

JEL classification: C31, O18, R11

[1] The authors acknowledge financial support from the Japanese Society for the Promotion of Science (JSPS). The corresponding author is Dr. V. Kravtsova, e-mail: vk2010@ier.hit-u.ac.jp. Postal address: Institute of Economic Research, 2-1 Naka, Kunitachi, Tel: +81-42-580-9146, Fax: +81-42-580-8333. We were able to gain access to the micro data of the *Census of Manufactures* as part of the research project on "Factory and Industry Level Analysis of Productivity" at the Research Institute of Economy, Trade and Industry (RIETI).

1. Introduction

A main cause of the long-term stagnation of the Japanese economy – the so-called two lost decades – appears to be the slowdown in total factor productivity (TFP) growth since the beginning of the 1990s. This slowdown in TFP growth has not only reduced Japan's potential GDP growth but also effective demand through stagnation in the rate of return to capital. Data – such as from the Japan Industrial Productivity (JIP) Database and the EU KLEMS Database – indicate that the slowdown in Japan's TFP growth from 1980s until the mid-1990s is mainly due to a drop in the pace of productivity growth in the manufacturing sector. This is illustrated in Figure 1, which shows that TFP growth in the manufacturing sector was on a downward trend until the first half of the 1990s. At the same time, TFP growth in the non-manufacturing sector has been very low for a long time.

Insert Figure 1

In order for Japan to accelerate TFP growth in the manufacturing sector the productivity of SMEs has to increase, while large and productive manufacturing firms need to expand their activities within Japan. These two issues are closely related: large firms improved their productivity by eliminating unnecessary factor inputs, affiliates, and supplier relationships (Fukao and Kwon 2006). The dissolution or weakening of supplier relationships may have reduced knowledge flows from large, productive final assemblers to smaller, less productive firms, which supply parts and components. Moreover, large firms did not expand production within Japan, partly because they relocated their factories abroad, which also contributed to weakening supplier relationships. In fact, as a consequence of the recent Great East Japan Earthquake and electricity shortages, acceleration in hollowing-out is expected. In a special survey on supply chains conducted by METI after the earthquake, 69% of the manufacturing firms surveyed answered that there is some likelihood that the whole or part of their supply chain will be relocated abroad because of the earthquake (Figure 2).

Insert Figure 2

The literature on knowledge flows shows that geographical proximity to research and development (R&D) activities and leading-edge firms frequently plays an essential role in knowledge spillovers (e.g., Jaffe, Trajtenberg, and Henderson 1993, Orlando 2004). Despite the importance of this issue, few studies have investigated where factories on the technology frontier are located and how proximity to these factories affects spillover effects in Japan. Against this background, the aim of this paper is to examine this issue using factory-level data of Japan's *Census of Manufactures*. We do so employing the following strategy. First, we estimate the efficiency of each factory by using a non-parametric data envelopment analysis (DEA) model for each industry. Second, we estimate the geographical distances to the most efficient factory in the prefecture and in all of Japan. Third, we identify the determinants of the performance of factories and estimate how geographical proximity to frontier factories affects spillover effects. We find that clustering occurs in each industry and efficient factories concentrate in certain regions. The percentage of efficient firms in total firms is particularly high in the Chubu and Tohoku regions. The estimation results also suggest that closeness to the most efficient factory plays a statistically significant role in determining the efficiency of manufacturing factories in Japan in most industries. However, this is not the case in high-tech industries.

The structure of the paper is as follows. The next section provides a review of the related literature. In Section 3, we then introduce our data and methodology. Next, Section 4, reports our estimation results, while Section 5 summarizes our main findings.

2. Literature Review

The various benefits of geographic concentration or agglomeration of economic activities, currently referred as clustering, have been widely discussed in the economic literature. The concept of agglomeration economies is attributed to Marshall (1890), Arrow (1962), and Romer (1986) (so called MAR spillovers)and is associated with industrial specialization and therefore an intra-industry phenomenon. Factories locate in close proximity to reduce the costs of purchasing from suppliers, or shipping to downstream customers.

More recently, the development of new trade theory (e.g., Krugman 1980, Krugman and Venables 1990) and new economic geography models (e.g., Krugman 1991, Krugman and Venables 1995, Fujita and Thisse 1996, Baldwin et al. 2003, and Fujita 1999) has resulted in "space" being recognized more widely as a crucial factor in determining economic development. The associated literature emphasizes the importance and role of knowledge assets in determining competitiveness, productivity, and ultimately output growth, by drawing a useful distinction between knowledge that is already internal to the factory (through learning-by-doing that draws on existing knowledge and human capital built up through R&D and similar investments) and knowledge gained externally (some of which is through market transactions, such as spending on extramural R&D, and some of which is gained through spillovers). Co-location, or reduction of geographical distance, is likely if there is a large, common pool of labor. Furthermore, knowledge spillovers occur when similar factories engage in innovative activities to solve similar or related problems involving both external and internal knowledge.

The approach usually taken in the spatial econometrics literature is to model geographical spillovers by determining the type and extent of spatial dependence that exists between areas by constructing spatial weights to reflect spatial interactions. Two types of spatial dependence are usually considered: the spatial lag (or autoregressive) model (equation 1) and a spatially weighted error term (or spatial error) model (equation 2). These two standard models are specified as follows:

$$y = \rho W\gamma + X\beta + u \tag{1}$$

$$y = X\beta + \boldsymbol{\lambda W \varepsilon} + u \tag{2}$$

where

***y*:** dependent variable,

X: matrix of independent variables with associated parameters β,

Wy: matrix of spatially lagged dependent observations,

$\boldsymbol{W\varepsilon}$: matrix of spatially lagged errors or a measurement error that is correlated with space

u: independent error term,

$\boldsymbol{\varepsilon}$: a spatially autoregressive error term,

$\boldsymbol{\rho, \lambda}$**:** parameters to be estimated that measure the strength of spatial autocorrelation in the model.

By requiring that spatial interaction be dealt with through the inclusion of other lagged variables in the model, the spatial lag model (equation 1) presumes that omission of *Wy* will result in omitted variable bias when estimating the parameters of interest (β). In contrast, the spatial error model (equation 2) treats spatial dependence as a statistical nuisance, assuming that such dependence occurs between variables that are not included in the model and which are therefore captured in ε. It has been argued (e.g., by Anselin 2003) that the researcher must determine which model best fits the data (i.e., whether ρ=0 or λ=0); however, a priori, the omission of variables from the model is undesirable because of the implications of misspecification (i.e., biased and inconsistent parameter estimates) and because, where the data permit, it is presumably more appropriate to treat spatial autocorrelation either by including additional relevant variables or by including spatially lagged values of the variables in the model to proxy for any missing variables. This point is often not discussed explicitly in the spatial econometrics literature (with rare exceptions like Andersson and Gråsjö 2009) and has implications for how the spatial weight matrix *W* is constructed.

To sort through a potentially large number of competing models it has become common practice to specify in advance a number of different versions of *W* and then use “goodness-of-fit” statistics to choose the model that best represents the data (see LeSage and Fischer 2008). As shown by Harris and Kravtsova (2009), following the practice above only local maxima among the competing models can be found, but not necessarily a correctly specified *W*.

Recent developments in estimating the role of distance can be seen as a result of the growing availability of data on the exact location of factories. Such data allow a precise estimation of the distance between factories and therefore more detailed specification of the model. The methodology developed in the current paper is inspired by, among others, the related work by Hanson (2005) and Harris (1954) using a market-potential function (see Appendix 1.2 for the methodological details). Namely, the performance of the factory is seen as a function of both internal and environmental factors, where a higher number of the most efficient factories in the prefecture and a shorter distance to them from other, less efficient factories in the prefecture are hypothesized to be a positive factor (more methodological details on how distance to the most efficient factory is measured and included in the regression analysis are provided in the next section of this paper).

Processes of knowledge generation and acquisition within the factory are essentially organizational learning processes (Reuber and Fisher 1997, Autio et al. 2000), and although factories could develop and acquire much of the knowledge internally (through their own resources and routines), few (and especially SMEs) rarely possess all the inputs required for successful and sustainable (technological) development. Therefore, meeting factories' knowledge requirements typically necessitates the use of external resources to acquire and internalize knowledge (Rosenkopf and Nerkar 2001, Almeida et al. 2003) and, as argued above, proximity is likely to be important when accessing such externalities.

The idea that close physical proximity (and density) play an important role is mainly predicated on the notion that a significant part of the knowledge that affects economic growth is tacit (and therefore difficult to codify). Such knowledge does not move readily from place to place as it is embedded in individuals and firms and the organizational systems of different places (Gertler 2003). This means that many kinds of spillovers are also limited by distance: the key channels for foreign direct investment (FDI) spillovers – labor turnover, demonstration effects, competition and cooperation with upstream suppliers (backward linkages) and downstream customers (forward linkages) – are geographically restricted in many industries.[2]

---

[2] For theoretical developments on the role of spatial agglomeration see the work of Liu and Fujita (1991), who use a monopolistic competition model to compare equilibrium urban configurations with optional configurations. Further, Fujita (2007) proposes the development of a "new economic geography," which

Another factor highlighted in the literature on spillovers is the decision to engage in export activity. To recoup the sunk costs of entry into overseas markets factories need to possess special knowledge assets that provide them with a comparative advantage (Dixit 1989, Baldwin and Krugman 1989, Roberts and Tybout 1997). Some empirical studies find that exporting is concentrated amongst a very small number of factories which nevertheless are large and account for the lion's share of trade (Bernard et al*.* 2005). It has also been confirmed that, compared with non-exporting indigenous factories, such exporters, *ceteris paribus*, have a greater probability of survival, much higher growth, are more productive and more capital-intensive, pay higher wages, and employ better technology and more skilled personnel (Eaton and Kortum 2001, 2002).

The level of competition between factories will also matter for spillovers. Incentives to learn from more efficient factories will clearly be strongest when the factories are in direct competition with each other and when passivity will result in lost market share and profits (Wang and Blomström 1992, Kokko 1996, Sjöholm 1999).

In a seminal article, Melitz (2003) extended Krugman's (1980) model to accommodate factory-level differences in productivity in order to analyze the intra-industry effects of trade. The model suggests that as a consequence of increasing exposure to trade, the most productive factories are stimulated to participate in export markets, while less productive factories continue to serve the domestic market only and the least productive factories drop out of the market. It follows that trade-induced reallocations towards more efficient factories will eventually lead to aggregate productivity gains. Other recent international trade models incorporating factory-level heterogeneity include the models by Bernard et al. (2003), which is based on Ricardian differences in technological efficiency, Helpman et al. (2004), which explicitly compares exporting and outward FDI as alternative modes of entry, Yeaple (2005), which focuses on heterogeneous competing technologies, trade costs, and labour skills, Bernard et al. (2007), which draws on heterogeneous productivity, and Aw et al. (2007, 2008, 2009), which add R&D as a new dimension to the export-productivity debate.

In other words, proximity to high-productivity factories is not the only determinant of potential productivity spillovers. Factories with foreign investment and those exposed to international trade (namely, exporters) posses asset-specific knowledge, which may potentially spill over to domestic factories. The characteristics that can make factories shoulder the (sunk) costs of entry into foreign

---

should be directed towards a comprehensive theory of spatial economics in the knowledge economy, where spatial economic dynamics are based on linkages between the fields of economics and knowledge.

markets and potentially have an impact on factories' profitability are size, labor composition, productivity, product mix, and ownership structure (Bernard and Jensen 2004). Bernard and Jensen (2004) find that other exogenous factors that can affect profitability are exchange rate movements, other shocks to demand, and indirect and direct subsidies to exporters and potential spillovers from the presence of other nearby exporters.

Yet another reason for differences in spillovers is that the behavior and strategies of factories may vary depending on their role in the corporate group to which this factory belongs. Given that the Japanese economy was outperforming that of the United States during the 1980s, some economists believed that the Japanese economic model, based on the use of very large horizontal and vertical conglomerates known as *keiretsu*[3], was superior to its American counterpart based on private market competition. The role of the *keiretsu* in Japan's productivity growth at that time has been examined by Miwa and Ramseyer (2006), who provide strong counter-arguments rejecting the significance of such structures of industrial organization as the *keiretsu*. Since the U.S. economy experienced the collapse of a financial bubble in 2008 similar to the burst of Japan's bubble in the early 1990s, the role of *keiretsu*-type organizational structure might be reevaluated. This study looks at the performance of factories taking two main types of corporate structure into account: the case when a factory is a part of a multi-factory group and the case when factory is a single-establishment one.

It has also been suggested that export-oriented factories may enjoy less scope for technology spillovers than import substituting local market-oriented affiliates (Javorcik 2004, Kokko et al. 2001). While local market-oriented factories typically bring with them technologies that are weak or missing in the host country, export-oriented affiliates are more likely to focus on activities and technologies where the host country already has comparative advantages. In this case, the competitive assets of the efficient factory may be superior marketing knowledge (related, for instance, to knowledge about competitors or access to existing distribution networks) rather than superior production technology. As a result, there is perhaps no reason to expect positive technology spillovers to other factories (although some of the knowledge related to exporting may well spill over).

Much of the discussion so far has been on how factories acquire and use knowledge, or what might be termed the "learning factory." In addition, factories show differential capabilities to absorb and translate available knowledge into (endogenous) economic growth. Maurseth and Verspagen

[3] These *keiretsu* were thought to follow the instructions of main banks and the Japanese government rather than their own entrepreneurial insights (Brennan 2008).

(1999: 152), for example, argue that the empirical evidence shows that the "ability to adapt new technologies depends on the institutional infrastructure, education, geography, and resources devoted to R&D." This highlights the importance of the *regional* innovation system in enabling factories to acquire external knowledge, i.e., the concept of the "learning region" (see, e.g., Cooke and Morgan 1998, Oughton et al. 2002, Cooke et al. 2003, Howells 2002, Asheim et al. 2005).

Based on the empirical literature on productivity, various studies have sought to examine the effect of regional development and externalities on the performance of firms in Japan. Otsuka et al. (2010), for example, use prefectural level data on spatial and industrial economic activities to assess the effect of externalities on the productive efficiency of Japanese regional industries. The study finds that agglomeration economies, defined as the presence of a concentration of firms belonging to the same industry in one location (economies of scale or MAR-spillovers), has a positive effect on the productive efficiency of manufacturing and non-manufacturing industries. In the current study, the availability of factory-level data allows us to take into account firm-level differences while looking at the cross-regional distribution of productivity in Japan. Later, the components of the regional system of innovation (proxied by the number of scientists, university graduates, etc., in the region) are taken into account to control for regional differences. A more detailed overview of the methodology used in the paper is presented in the next section.

3. Data and Methodological Background

The dataset used in this study has been constructed by merging factory-level data with prefecture- and industry-level data. The factory-level data are based on the 2007 *Kogyo Tokei Chosa* (*Census of Manufactures*), which is conducted annually by the Ministry of Economy, Trade and Industry (METI) and covers the economic activities in 2007 of all Japanese manufacturing factories except those belonging to the government and offices not directly engaged in manufacturing, processing, or the repair of industrial products. The 2007 census data cover factories with four or more employees and exclude small factories due to the lack of information on capital stock. We merged the factory data with data from the *Population Census* on the estimated population density of the prefecture, prefectural local government R&D, the share of university graduates in the prefectural population, the share of managerial and technical employees in total prefectural employment, and the propensity to engage in export activity at the industry and prefecture level. In order to apply OECD technology-based industry classifications to the factory data, we converted the Japan Standard Industry Classification to the International Standard Industry Classification using the industry conversion table developed by METI. In the next step, we then classified industries into high-tech, medium high-tech, medium low-tech, and low-tech industries (see Hatzichronoglou 1997 for details on the classification by level of technology).

To estimate the relevance of the main determinants of the performance of factories, we proceed in three steps: first, we estimate factory-level efficiency using DEA analysis; second, we estimate the geographical distance from the most efficient to the least efficient factory; and third, we perform regression analysis to identify the determinants of factories' performance.

First Step: Efficiency Estimation

To estimate the efficiency of factories using data envelopment analysis (DEA), the parameters of the production function were specified as follows: output – shipments; inputs – materials, capital stock and wages. Prices on inputs are assumed not to vary greatly among industries within Japan. Therefore, the technology available to a factory at a given point in time (2007) defines which input-output combination is feasible. It is assumed that factories can maximize their output for a given amount of inputs they have. In the absence of market prices, DEA endogenously generates "shadow prices" of inputs and output for aggregation (see Appendix 1.1 for the technical details of the DEA analysis).

The deterministic assumption used in DEA models that all observed units belong to the attainable set requires a robust procedure for outlier detection (Simar 2003). Since envelopment estimators are very sensitive to extreme observations they can behave dramatically in the presence of super-efficient factories, which can be viewed as outliers. The exploratory data analysis procedure recently proposed by Simar (2003), which is more robust to the presence of observations on super-efficient factories, was used. Employing this procedure, no outliers were detected in the sample. An efficiency score was estimated for each factory j out of the sample of n factories independently for each industry, while keeping a common frontier across all prefectures in Japan. Therefore, 52 models, one for each 2-digit industry, where factories share a common technology frontier, are estimated. Firms that score the maximum of 1 (unity) in each industry and form the technological frontier are deemed to be the most efficient.

We then compare the distribution of the efficiency scores of all factories across all sectors and all prefectures with the normal distribution. Estimated DEA efficiency scores tend to have a bimodal distribution (the distribution shows two peaks, at about 0.25 and 0.95, as can be seen in Figures 7 to 9). This suggests that there are relatively large numbers of factories with very low efficiency and with high efficiency in our sample.

Second Step: Geographical Distance Estimation

In the second step, we calculate the distance to the most efficient factory (see Appendix 1.2 for the technical details on calculating bilateral distances). The full detailed address of each factory has been converted to the longitude and latitude geographical location of the factory, and we use great-circle distances to calculate the factory-unique distance to the most efficient factory in the prefecture.[4] This distance is used later to estimate spillover potential based on geographical closeness to the most efficient factory, using the following equation proposed by Harris (1954):

$$SP_j \equiv \sum_{k \in E} \frac{1}{d_{jk}} \tag{3}$$

where E is the set of efficient (exporting) factories, and $d_{j,k}$ is the bilateral distance between factories j and k.

The spillover effect from efficient (exporting) factories might decay more quickly as distance increases than equation (3) implies. Thus, according to Hanson's (2005) definition of market potential, spillover potential is alternatively defined as follows:

$$SP_j \equiv \sum_{k \in E} \exp(-d_{jk}) \tag{4}$$

This allows a more accelerating decay of the spillover effect as distance increases than implied in equation (3). In order to check the robustness of our results, we use both definitions and report the results in the tables.

Third Step: Determinants of Efficiency and Distance

In the third step, the efficiency score obtained from the DEA analysis described in the previous section is regressed on environmental variables. The purpose of this step is to account for exogenous factors (e.g., industry- or prefecture-specific factors) that might affect factories' performance and cannot be directly taken into account in the first-step non-parametric model. The general model for the second stage can be specified as follows:

$$\delta_j^* = Z_j \beta + \tau_j \tag{5}$$

where δ_j* indicates the estimated technical efficiency score of each factory j. Since the estimates are

[4] In order to check the robustness of the results, the distance to the most efficient factory within all Japan, without taking prefectural boundaries into account, was re-estimated. Both types of distance calculation provide similar results (see Tables 3 and 4).

bounded by unity in output-oriented models, it has been argued that DEA efficiency estimates are truncated. In order to take the truncation problem into account in a coherent manner, Simar and Wilson (2007) propose an approach based on truncated regression where the error term τ_j is identically and independently distributed for all j with $N(0, \sigma_\varepsilon^2)$. Further, they point out that the conventional approaches to inference employed in many studies, which rely on multi-stage approaches, are invalid due to complicated unknown, serial correlation among the estimated efficiencies. The criticism applies equally to the use of a "naïve" bootstrap in Hirschberg and Lloyd (2002).

Following Simar and Wilson's (2007) algorithm 1 procedure, we use the maximum likelihood method to obtain the estimate $\hat{\beta}$ of β as well as estimates $\hat{\sigma}_\varepsilon$ of σ_ε in the truncated regression of equation (6). The bootstrap estimates were obtained by following the three steps in Simar and Wilson (2007) and the confidence interval was defined based on bootstrapped values of β and σ_ε. The more detailed empirical model looks as follows:

$$E_{j,p,i} = f\,(\text{Factory\_Characteristics}_{j,p,i}, \text{Industry\_Characteristics}_{i}, \text{Prefecture\_Characteristics}_{p}) + \tau_{j,p,i} \qquad (6)$$

where j: factory, i: industry, p: prefecture, $\tau_{j,p,i}$: error term, and $E_{j,p,i}$: the factory's estimated technical efficiency score. Here, the frontier was estimated for each industry separately for a production possibility set which initially contains observations of all types of factories (see previous section for details). In this stage, only the technical efficiency of factories that scored less than 1 is used for the dependent variable in order to capture the spillover potential based on the geographical closeness to the most efficient factory.

4. Results Analysis

The distribution of manufacturing factories in Japan is presented in Figures 4 to 6.

We start by looking at the distribution of factories across broad regions. Specifically, we divide Japan into the following eight regions, each of which consists of a number of prefectures (except the Hokkaido region, which consists only of Hokkaido):

A. Hokkaido (the island of Hokkaido and nearby islands, largest city: Sapporo)
B. Tohoku region (northern Honshu, largest city: Sendai. Prefectures in this region were most severely hit by the recent earthquake and tsunami on March 11, 2011)

C. Kanto region (eastern Honshu, largest city: Tokyo)
D. Chubu region (central Honshu, including Mt. Fuji), sometimes divided into:
E. Hokuriku region (northwestern Chubu, largest city: Kanazawa)
F. Koshinetsu region (northeastern Chubu, largest city: Niigata)
G. Tokai region (southern Chubu, largest city: Nagoya)
H. Kinki region (west-central Honshu, largest city: Osaka)
I. Chugoku region (western Honshu, largest city: Hiroshima)
J. Shikoku (island, largest city: Matsuyama)
K. Kyushu (island, largest city: Fukuoka), which in our regional division includes the Ryukyu Islands, including Okinawa

Insert Figure 3

Figures 3(a) to (d) show the distribution of factories at the aggregate regional level and allow us to identify the pattern of distribution of all and the most efficient factories. Figures 3(a) and 3(b) suggest that the absolute number of factories and the number of efficient factories in a region are highly correlated. Namely, regions in the center of Japan, like the Kanto and Chubu regions, host both the highest number of manufacturing factories and the highest number of efficient factories, while the Southern Chugoku and Shikoku regions and Hokkaido have the lowest absolute number of factories and also the lowest concentration of efficient factories. To some extent, the pattern shown in Figure 3(b) simply reflects the fact that manufacturing activity in certain regions is sparse. Therefore, in Figure 3(c), we show the number of efficient factories relative to the total number of factories in a region. This indicates that the clear leadership of the central region of Japan diminishes, with Kanto no longer at the top, and that Tohoku is now on par with Chubu. At the same time, the position of regions like Chugoku, Shikoku, and Hokkaido remains unchanged, indicating that in relative terms, too, these regions have very few efficient factories. Another "redistribution" takes place in Figure 3(d), where the number of efficient factories is normalized by the area of a prefecture. Here, the Kanto region regains its leadership position and the relatively small Kinki region has a relatively high level of efficient factories.

The distribution of the number of factories by prefecture is mapped in Figure 4(a). This figure suggests that, apart from Tokyo and the surrounding prefectures, the vast majority of manufacturing factories is located not in the Kinki area (comprising Osaka, Kobe and Kyoto), but in Aichi prefecture. Aichi is known to be a cluster for manufacturing factories of such big Japanese corporations as Toyota, Fuji Heavy Industries, Denso, Mitsubishi Motors, Sony, and Suzuki, and of affiliates of foreign firms such as Bodycote (U.K.) and Pfizer (U.S.A.).

Insert Figure 4

Other prefectures accounting for a relatively large number of factories include Saitama (host of Honda Corporation and a number of factories in the food, optical, precision, biotechnology, and pharmaceutical industries), Kanagawa (which has a strong economic base in the shipping, biotechnology, and semiconductor industries), Tokyo (although it is a hub for corporate headquarters and service industries, it also hosts a range of factories of multinationals such as Fujitsu, Toshiba, and NEC), Shizuoka (which hosts factories in advanced health-related industries such as pharmaceuticals, medical equipment, etc., in food and chemical products-related industries, and in optical and electronic technology-related industries), Hyogo (which has many factories in heavy industries, metal and medical instruments) and Osaka (which hosts globally renowned electronics giants such as Hitachi Maxell, Sharp, Panasonic, and Sanyo, which was recently acquired by Panasonic, as well as factories in other industries).

Figure 4(b) shows the distribution of the most efficient factories (factories that are on the frontier or score unity). The prefectures with highest number of the most efficient factories are Kanagawa, Aichi, and Hyogo. These results suggest that there is clustering taking place and that the concentration of most efficient factories in prefectures with the highest density of manufacturing factories may be the result of productivity spillovers. To test the hypothesis of clustering or high concentration of manufacturing factories in one prefecture, we use further regression analysis in the second stage. Meanwhile, it is notable that prefectures such as Hokkaido, Tochigi, Saitama, Tokyo, Shizuoka and Fukuoka follow the leading group of prefectures in terms of the number the most efficient factories.

The visual difference between Figures 4(a) and (b) is that some remote prefectures like Fukuoka and Hokkaido, with a relativity low density of manufacturing factories, still have a high number of efficient manufacturers.

In the next step, we divide the sample of all efficient factories into four groups according to the OECD industry classification based on the technological intensity. The distribution of efficient factories within each of the four groups is mapped in Figure 5 and suggests that the distribution of efficient factories within the four groups of manufacturing industries differs. Namely, high-tech manufacturing (Figure 5(a)) shows particularly strong geographical clustering in the central prefectures of Japan, such as Aichi, Shizuoka, Kanagawa and Saitama as well as Tokyo, Hyogo and Tochigi, while medium high-tech (Figure 5(b)) and medium low-tech (Figure 5(c)) manufacturing show a relatively even geographical distribution of efficient manufacturing factories. On the other hand, low-tech manufacturing (Figure 5(d)) shows a relatively high concentration of efficient factories in remote areas such as Hokkaido and Fukuoka as well as in central parts of Japan, pointing to the presence of interesting underlying economic factors affecting the performance of factories in these prefectures. Thus, the overall pattern that emerges is that while Aichi prefecture plays host to the most efficient factories in all four types of industries (Figures 5(a) to (d)), remote Hokkaido, for example, scores relatively highly only in low-tech sectors.

These preliminary results indicate a certain prefectural specialization in Japan. However, they say very little about the significance of the role of distance in the performance of factories operating in Japan. Therefore, in the next step we conduct a regression analysis in which we examine the role of distance in factory performance while controlling for other environmental parameters such as factory, industry, and prefectural characteristics

Although there continues to be a discussion in the literature regarding whether the fact that a factory engages in multinational activity is an endogenous or exogenous determinant of its performance, what is beyond doubt is that factors related to multinational activity such as exports are important. Having built a common frontier for all factories in each sector separately, we can plot the kernel density distribution of the efficiency scores of exporting and non-exporting factories in all manufacturing sectors taken together. As shown in Figure 6, doing so suggests that exporters perform better than non-exporters, but the margin is not very large.

Insert Figure 6

One of the reasons why the margin is not very large is that the domestic market in Japan in many industries is close to the global frontier and the additional knowledge that factories engaged in export activity can theoretically bring back to Japan is relatively limited. In other words, the productivity externalities from export activities in Japan are not very high. On the other hand, as shown in Figure 7, there are clear differences in the productivity distributions of large factories and small and medium factories (SMEs). Large factories show a higher concentration at the efficient end than SMEs and a lower concentration at less efficient levels than in the case of SMEs. This indicates that in Japan size plays an important role in the performance of factories. One possible explanation of this difference is that larger firms may form part of *keiretsu* or network arrangements. In order to examine this issue, Figure 8 plots the kernel density distributions for single-factory establishments and for factories that form part of multi-factory establishments, i.e., factories that form part of a group of factories or a *keiretsu*. However, although factories belonging to multi-factory establishments show a slightly better performance, the figure does not provide clear evidence that this factor plays a major role in explaining the difference between large factories, which are more likely to be part of multi-factory establishments, and SMEs.

Insert Figures 7 and 8

To investigate the role and significance of such factory-level factors as size and whether the factory forms part of a multi-factory firm or is a single-factory establishment, as well as various industry and regional characteristics, we conduct further regression analysis. The results are presented in Table 1. Specifically, we estimate six models using Harris's (1954) definition of physical (geographical) distance and apply it to estimate the distance to the most efficient factories with a score equal to unity (Models 1-6). To check the robustness of the results, we then re-estimate the six models using another definition of physical distance (Hanson 2005) and apply it to estimate the distance to the most efficient factories (Models 7-12). The main difference between the two definitions can be summarized as follows: while in case of Harris's (1954) model physical distance is the sum of inverse distances to the most efficient factories (establishments that form a technological frontier and score highest on the efficiency scale), in the case of Hanson's (2005) model physical distance is the exponential (accelerating) way of distance decay. In all twelve models the dependent variable is the efficiency score obtained from the DEA estimation using variable returns to scale (VRS).

Insert Table 1

Looking at the role of distance in the performance of factories in all twelve models in Table 1, a positive and significant effect of closeness to the most efficient factories can be identified. This result implies that the closer manufacturing factories are to their most efficient counterparts, the better they perform. While these results can arguably imply the endogenous clustering of factories in one location, the importance of this effect on the performance of the factories in Japan is strongly supported by the results presented in Table 1.

Let us now examine the results in greater detail. The baseline OLS specification (Model 1) shows a significant positive correlation between the efficiency of a factory and the distance to the most efficient factories in each prefecture. Next, Model 2 re-estimates this basic model using truncated regression. The results show that the correlation becomes even stronger. In the following models (Models 3 to 6), we introduce various controls for factory-level, industry, and prefectural heterogeneity.

The results for Models 3 and 4 suggest that bigger factories tend to be more efficient than their smaller counterparts. Since we estimated our dependent variable under the assumption of variable returns to scale, this outcome suggests that scale efficiency is present in the manufacturing sector. Furthermore, the regression analysis confirms our previous observation that exporters are more efficient than factories not engaged in export activity. While the margin varies around the relatively small value of 0.02, it remains highly significant in all specifications. The coefficient estimates for the multi-factory dummy suggest that factories, which are part of a multi-factory firm, tend to be more efficient. A likely reason for this outcome is that such factories enjoy synergy effects, where operations such as accounting and distribution are largely centralized and the factories therefore do not directly bear the costs for these activities.

Next, let us look at the results for the variables representing prefectural characteristics, such as population density, R&D expenditures harmonized by population, and the share of manager employment in total employment. The coefficients on these variables tend to be negative, although only in the case of R&D expenditure are they consistently significant. On the other hand, the coefficient on the share of technician employment is positive and significant in all specifications where it is included, indicating that there is a clear correlation between the share of skilled technical workers in the population and factory efficiency in the prefecture.

Models 3 to 6 also include prefecture-level dummies (omitted in the table) and technology-intensity dummies to characterize industries. The results for the technology-intensity dummies suggest that factories in the less technology-intensive industries tend to be more efficient than their counterparts in the high-tech industries.

In Model 6, the importance of exporting is examined by including variables on the share of exports in sales in the prefecture and in the industry. The results suggest that in prefectures and industries with a higher share of exports in total sales factories perform worse than their counterparts in a less competitive environment. Recalling the distribution shown in Figure 6, which indicated that exporters are more efficient than non-exporters, the result obtained in Model 6 indicates negative intra- and inter-industry export spillovers. One possible reason is the increased competition within the industry and prefecture that non-exporting factories face in the presence of big exporters, forcing non-exporters downs the technological ladder. To examine this hypothesis, in Table 2 below we estimate the distance to exporters and run regressions separately for factories in the high-tech, high medium-tech, low medium-tech and low-tech industries. Before that however, in order to check the robustness of our results, we re-estimated all six models using the alternative definition of physical distance (Models 7 to 12). The results are very similar to those for Models 1 to 6, suggesting that the results are robust to our definition of physical distance.

We now turn to the examination of the role of distance to exporters. Specifically, we compare the effect on efficiency of distance to the most efficient factories and distance to exporters. The first four models shown in Table 2 suggest that the distance to the most efficient factory plays a positive role in the performance of other factories in all industries apart from the high-tech industry. This finding potentially suggests that "distance decay" is greater the higher the technology-level of an industry. That is, for factories in high-tech industries, the presence of other efficient factories does not play a significant role in determining their performance. At the same time, when we look at the role of the distance to exporters in the prefecture (Models 5 to 8), this distance tends to be important only for factories in low-tech industries. One possible explanation is that in these industries, exporters' knowledge on higher international quality standards may spill over to nearby suppliers and customers and thereby raise their efficiency. Finally, consistent with Figure 4, in all four types of industries exporters are more efficient than non-exporters counterparts. Factory size also continues to be associated with greater efficiency in all regressions.

Insert Table 2

Finally, in Tables 3 and 4, we show the regression results when distances are measured without prefecture border restrictions, i.e., we measure the distance not to the most efficient factory within the same prefecture, but to the most efficient factories within all of Japan. The signs on the coefficient do not change in any of the regressions, while the significance of the distance drops in some specifications (such as in Model 4 in Table 3, for example). In Table 4, some of the regressions (Models 3 and 7) show an increased significance of distance, once this is measured as the distance to the most efficient factory in Japan as a whole. The results suggest that the distance to the most efficient factory and to exporters within Japan is more important for low medium-tech and low-tech factories than for medium high-tech and high-tech factories.

Insert Tables 3 and 4

5. Conclusion

Using an original factory-level dataset, this paper provides a unique analysis of the role of distance between factories and their performance. Using non-parametric techniques, data points on three inputs and one output have been enveloped by a technological frontier (using a DEA procedure). Identifying the most efficient factories (those on the frontier) and their location, we then examined the distribution patterns of factories. We found that the absolute number of factories and the number of efficient factories in a region are highly correlated. Specifically, central regions of Japan host the highest number of manufacturing factories, while southern and northern regions have the lowest absolute number of factories. Even though the vast majority of manufacturers are located in the central part of the country, the leading “host prefecture” is not Tokyo or Osaka, but Aichi. The prefectures with the highest number of the most efficient factories are Kanagawa, Aichi and Hyogo. These results suggest that there is a clustering effect taking place and that the concentration of most efficient factories in prefectures with the highest density of manufacturing factories may be the result of productivity spillovers. To test the hypothesis of clustering of manufacturing factories in one prefecture, regression analysis was used in the next stage.

Moreover, the sample of all efficient factories was divided into four groups in accordance with OECD industry classifications based on the technological intensity. The results for the Kernel density distribution for the four sub-samples suggest that within the groups of manufacturing industries, the distribution of efficient factories differs. Namely, high-tech industries show particularly strong geographical clustering in the central prefectures of Japan, while medium high-tech and medium low-tech industries show a relatively even geographical distribution of efficient factories. In contrast, the relatively high presence of low-tech efficient factories in remote areas of Japan as well as in the central part of Japan reveals the presence of interesting economic factors underlying the performance of factories in these prefectures.

These preliminary results indicated a certain prefectural specialization in Japan, but said little regarding the significance of the role of distance in the performance of factories operating in Japan. We therefore conducted a regression analysis to examine the role of distance to the most efficient factory in factories’ performance while controlling for other environmental parameters such as factory, industry, and prefectural characteristics. The empirical results imply that the closer factories are to their most efficient counterparts, the better they perform. That is, the factory-level data on factories operating in Japan confirm that geographic proximity plays a positive and significant role in determining factories' efficiency. While the endogeneity issue behind the clustering of factories in one location can arguably drive this outcome, the significance of this phenomenon is strongly supported by the results. At the same time, the results also imply that geographical proximity is not an important determinant of the efficiency of factories in high-tech industries. This may be the result of a phenomenon known as "distance decay;" that is, for factories in Japan's high-tech industries, proximity to efficient factories does not matter.

Another finding was that exporters tend to be more efficient than non-exporters, although the difference was found to be not very great. One possible reason why the difference is not very great is that many industries in the domestic market in Japan are close to the global frontier and that therefore the additional knowledge that factories engaged in export activity can theoretically bring back to Japan is relatively limited. In other words, productivity externalities from export activities in Japan are not high. To further examine the role of exporting and other factors that can affect the performance of factories, we carried out additional regression analyses. We found that, in the case of low-tech industries, geographical proximity to exporting factories tends to go hand-in-hand with higher efficiency, suggesting that in these industries clustering around export-oriented and efficient factories is beneficial for factories. At the same time, a high concentration of export activities at the prefectural level appears to have negative spillover effects, possibly reflecting the effect of competition in a particular region or prefecture. These results suggest that from a factory's perspective, it is beneficial to be located close to efficient and export-oriented front-runners, while at the same time from a regional policy perspective, a high concentration of export activity within one prefecture might have a negative effect on the performance of other factories in the prefecture.

References

Almeida, P., Dokko, G., and Rosenkopf, L. (2003) "Startup Size and the Mechanisms of External Learning: Increasing Opportunity and Decreasing Ability?" *Research Policy*, 32(2), 301-315.

Andersson, M., and Gråsjö, U. (2009) "Spatial Dependence and the Representation of Space in Empirical Models," *The Annals of Regional Science*, 43(1), 159-180.

Anselin, L. (2003) "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics," *International Regional Science Review*, 26(2), 153-166.

Arrow, K. (1962) "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29(2), 155-173.

Asheim, B.T., Gertler, M.S., Fagerberg, J., Mowery, D.C., and Nelson, R.R. (2005) "The Geography of Innovation: Regional Innovation Systems," in *The Oxford Handbook of Innovation*, Oxford and New York: Oxford University Press.

Autio, E., Sapienza, H. J., and Almeida, J. G. (2000) "Effects of Age at Entry, Knowledge Intensity, and Imitability on International Growth," *The Academy of Management Journal*, 43(5), 909-924.

Aw, B.Y., Roberts, M.J., and Winston, T. (2007) "Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity," *The World Economy*, 143, 83-104.

Aw, B.Y., Roberts, M.J., and Xu, D.Y. (2008) "R&D Investments, Exporting, and the Evolution of Firm Productivity," *American Economic Review*, 98, 451-456.

Aw, B.Y., Roberts, M.J., and Xu, D.Y. (2009) "R&D Investment, Exporting, and Productivity Dynamics," *NBER Working Paper* 14670. National Bureau of Economic Research.

Banker, R.D., Charnes, A., and Cooper, W.W. (1984) "Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis," *Management Science*, 30, 1078-92.

Baldwin, R. and Krugman, P. (1989) "Persistent Trade Effects of Large Exchange Rate Shocks," *Quarterly Journal of Economics*, 104, 635-654.

Baldwin, R., and Martin, P. (2003) "Agglomeration and Regional Growth," *C.E.P.R. Discussion Papers* 3960, CEPR.

Bernard, A.B., and Jensen, J.B. (2004) "Entry, Expansion, and Intensity in the US Export Boom, 1987-1992," *Review of International Economics*, 12(4), 662-675.

Bernard, A.B., Eaton, J., Jensen, J.B., and Kortum, S. (2003) "Factories and Productivity in International Trade," *American Economic Review*, 93, 1268-1290.

Bernard, A.B., Jensen, J.B., and Schott, P.K. (2005) "Importers, Exporters, and Multinationals: A

Portrait of Firms in the U.S. that Trade Goods," *NBER Working Paper* 11404, National Bureau of Economic Research.

Bernard, A.B., Redding, S., and Schott, P.K. (2007) "Comparative Advantage and Heterogeneous Firms," *Review of Economic Studies*, 74(1), 31-66.

Brennan, T.J. (2008) "Conspiracy in Industry or Conspiracy in Academe? A Review of Yoshiro Miwa and J. Mark Ramseyer, The Fable of the Keiretsu: Urban Legends of the Japanese Economy," *International Journal of the Economics of Business*, 15, 265-79.

Coelli, T. (1996) "A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program," *Centre for Efficiency and Productivity Analysis (CEPA) Working Paper* 08, CEPA.

Coelli, T.J., Rao, D.S.P. and Battese, G.E. (1998) *An Introduction to Efficiency and Productivity Analysis*, Boston and Dordrecht: Kluwer Academic Publishers.

Cooke, P.N., and Morgan, K. (1998) *The Associational Economy: Firms, Regions, and Innovation*, Oxford and New York: Oxford University Press.

Cooke, P., Roper, S., and Wylie, P. (2003) "The Golden Thread of Innovation and Northern Ireland's Evolving Regional Innovation System," *Regional Studies*, 37(4), 365-379.

Dixit, A. (1989) "Entry and Exit Decisions under Uncertainty," *Journal of Political Economy*, 97, 620-638.

Eaton, J., and Kortum, S. (2001) "Technology, Trade, and Growth: A Unified Framework," *European Economic Review*, 45(4-6), 742-755.

Eaton, J., and Kortum, S. (2002) "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741-1779.

Färe, R. and Primont, D. (1995) *Multi-output Production and Duality: Theory and Applications*, Boston and Dordrecht: Kluwer Academic Publishers.

Fujita, M., Krugman, P. and Venables, A., 1999, *The Spatial Economy: Cities, Regions, and International Trade*: MIT.

Fujita, M., and Thisse, J. F., 1996, "Economics of Agglomeration," *Journal of the Japanese and International Economies,* 10(1), 339-78.

Fujita, M. (2007) "Towards the New Economic Geography in the Brain Power Society," *Regional Science and Urban Economics*, 37, 482-90.

Fukao, K., and Kwon, H.U. (2006) "Why Did Japan's TFP Growth Slow Down in the Lost Decade? An Empirical Analysis Based on Firm-Level Data of Manufacturing Firms," *Japanese Economic Revie*w, 57(2), 195–228.

Fukao, K., and Kwon, H.U. (2010) "Nihon no Keizai Saisei no Gendoryoku wo Motomete (The Key

Drivers of Future Growth in Japan)," Paper prepared for the Growth Strategy Research Project of the American Chamber of Commerce in Japan (ACCJ), ACCJ.

Hanson, G.H. (2005) "Market Potential, Increasing Returns and Geographic Concentration," *Journal of International Economics*, 67, 1–24.

Harris, C.D. (1954) "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers*, 44, 315-48.

Hamaaki, J., Hori, M., Maeda, S., and Murata, K., (2001) "Changes in the Japanese Employment System in the Two Lost Decades," *PIE/CIS Discussion Paper Series* 511, Center for Intergenerational Studies, Institute of Economic Research, Hitotsubashi University.

Harris, R. and Kravtsova, V. (2009) "In search of 'W'," *SERC Discussion Papers* SERCDP0017, Spatial Economics Research Centre (SERC), London School of Economics and Political Science, London, UK.

Hatzichronoglou, T. (1997) "Revision of the High-Technology Sector and Product Classification," *STI Working Papers*, 1997/2 OECD, Paris.

Hayashi, F., and Prescott, E.C. (2002) "The 1990s in Japan: A Lost Decade," *Review of Economic* Dynamics, 5(1), 206-35.

Helpman, E., Melitz, M.J., and Yeaple, S.R. (2004) "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 94, 300-316.

Howells, J.R.L. (2002) "Tacit Knowledge, Innovation and Economic Geography," *Urban Studies*, 39, 871-84.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993) "Geographical Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics* 108, 577-598.

Javorcik, S.B. (2004) "The Composition of Foreign Direct Investment and Protection of Intellectual Property Rights: Evidence from Transition Economies," *European Economic Review*, 48, 39-62.

Kokko, A. (1996) "Productivity Spillovers from Competition between Local Firms and Foreign Affiliates," *Journal of International Development*, 8, 517-30.

Kokko, A., Zejan, M. and Tansini, R. (2001) "Trade Regimes and Spillover Effects of FDI: Evidence from Uruguay," *Weltwirtschaftliches Archiv*, 137, 124-49.

Kravtsova, V. (2008) "Foreign Presence and Efficiency in Transition Economies," *Journal of Productivity Analysis*, 29, 91-102.

Krugman, P. (1980) "Scale Economies, Product Differentiation, and the Pattern of Trade," *American*

Economic Review, 70(5), 950-959.

Krugman, P. (1991) *Geography and Trade*, Cambridge, MA: MIT Press.

Krugman, P.R., and Venables, A.J. (1995) "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110 (4), 857-880.

Krugman, P.R., and Venables, A.J. (1990) "Integration and the Competitiveness of Peripheral Industry," *CEPR Discussion Papers* 363, CEPR.

Krugman, P. (1998) "Japan's Trap," online: http//web.mit.edu/krugman/www/japtrap.htm.

Liu, H.-L. and Fujita, M. (1991) "A Monopolistic Competition Model of Spatial Agglomeration with Variable Density," *The Annals of Regional Science*, 25, 81-99.

LeSage, J.P., and Fischer, M.M. (2008) "Spatial Growth Regressions: Model Specification, Estimation and Interpretation," *Spatial Economic Analysis*, 3(3), 275-305.

Marshall, A. (1890) *Principles of Economics* (revised edition), London: Macmillan.

Maurseth, P.B., and Verspagen, B. (1999) "Europe: One or Several Systems of Innovations? An Analysis Based on Patent Citations," in J. Fagerberg, P. Guerrieri, and B. Verspagen (eds.), *The Economic Challenge for Europe: Adapting to Innovation Based Growth*, Cheltenham, U.K. and Northampton, Mass.: Elgar.

Melitz, M.J. (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695-725.

Miwa, Y., and Ramseyer, J.M. (2006) *The Fable of the Keiretsu: Urban Legends of the Japanese Economy*, Chicago: University of Chicago Press.

Orlando, M. (2004) "Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity," *Rand Journal of Economics*, 35, 777-786.

Otsuka, A., Goto, M., and Sueyoshi, T., (2010) "Industrial Agglomeration Effects in Japan: Productive Efficiency, Market Access, and Public Fiscal Transfer," *Papers in Regional Science*, 89 (4), 819-40.

Oughton, C., Landabaso, M., and Morgan, K. (2002) "The Regional Innovation Paradox: Innovation Policy and Industrial Policy," *Journal of Technology Transfer*, 27 (1), 97-110.

Reuber, A.R., and Fischer, E. (1997) "The Influence of the Management Team's International Experience on the Internationalization Behaviors of SMEs," *Journal of International Business Studies*, 28(4), 807-825.

Roberts, M.J., and Tybout, J.R. (1997) "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87 (4), 545-564.

Romer, P.M. (1986) "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94, 1002-37.

Rosenkopf, L., and Nerkar, A. (2001) "Beyond Local Search: Boundary-Spanning, Exploration, and Impact in the Optical Disk Industry," *Strategic Management Journal*, 22 (4), 287-306.

Shephard, R.W. (1970) *Theory of Cost and Production Functions*, Princeton: Princeton University Press.

Simar, L., and Wilson, P.W. (2007) "Estimation and Inference in Two-stage, Semi-parametric Models of Production Processes," *Journal of Econometrics*, 136(1), 31-64.

Sjöholm, F. (1999) "Productivity Growth in Indonesia: The Role of Regional Characteristics and Direct Foreign Investment," *Economic Development and Cultural Change*, 47, 559-84.

Wang, J.-Y. and Blomström, M. (1992) "Foreign Investment and Technology Transfer: A Simple Model," *European Economic Review*, 36, 137-55.

Yeaple, S.R. (2005) "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics*, 65, 1-20.

Appendix 1: Economic and Physical Distances in the Derivation of Efficiency Spillovers

1.1. Economic distance

In this study, the performance of the economic units (factories in our case) is measured using a deterministic (non-parametric) approach. In contrast to traditional methods, no a priori assumption on the functional form of the frontier are made prior to the estimation and instead observations on inputs and output are put in the hyperspace and enveloped by the technological frontier using data envelopment analysis (DEA). The frontier has been build for each industry of the Japan Industrial Productivity (JIP) database classification (for details of the classification, see Fukao 2008). The frontier represents the existing technology that is formed by the most efficient factories in a particular industry. Using the cost-minimizing approach, these ”front runners” get a score equal to unity and the remaining factories, which are less efficient and therefore behind the frontier, score always below unity and get a score according to the estimated distance to the frontier. Therefore, the further a factory is from the frontier the less efficient the factory is and the lower its score

To facilitate the formal discussion let $x_i^j = (x_{i,1}^j,....,x_{i,N}^j)' \in \Re^N$ be a vector of N inputs that each factory j (j = 1, 2,... , n) uses to produce a vector of M outputs, denoted by $y_i^j = (y_{i,1}^j,....,y_{i,N}^j)' \in \Re^M$.

We assume that the technology of any factory j in industry i is characterized by the output set

$$E_i^j(x^j)=\{y:\ y \in \Re^M \text{ is producible from } x^j \in \Re^N\} \qquad \text{(A1)} –$$

The technology in any industry i satisfies the usual regularity axioms of production theory (see, e.g., Färe and Primont 1995), so that we can use Shephard’s (1970) *output oriented* distance function D_i^j: $\Re^N \times \Re^M \to \Re^1 \cup \{\infty\}$, defined as

$$D_i^j(x^j, y^j)=\inf\{\theta: y^j \mid \theta \in E_i^j(x^j)\} \qquad \text{(A2)}$$

In the choice of orientation of the model (input vs. output-oriented), we chose the output orientation, reflecting the assumption that in manufacturing sectors factories have more control over outputs than inputs. At the same time, the choice of appropriate orientation of the model is not as crucial as in the case of parametric analyses and reflects the choice of output maximization or input minimization in the linear programming model.

The linear programming model for the output-oriented variable returns to scale (VRS) model is:

$$\max_{\theta,\lambda}\ \theta,$$

subject to

$$-\theta y_i+Y\lambda\geq 0,$$
$$x_i-X\lambda\geq 0, \qquad \text{(A3)}$$
$$N1'\lambda=1$$

$\lambda \geq 0$,

where $1 \leq \theta < \propto$, and θ–1 is the proportional increase in outputs that will be taken for the i-th decision making unit by holding input quantities constant. Y and X are output and input, respectively. N*1 is an N*1 vector of one. Note that 1/θ defines the output-orientated Variable Returns to Scale Technical Efficiency score reported as VRSTE in the paper (as in Coelli 1996).

The estimates of the technical efficiency score indicate the extent to which it is possible for a factory to increase output with input quantities held constant. This model incorporates a dual approach with a correction for slack (Coelli et al. 1998, Coelli 1996) and VRS, as suggested by Banker et al. (1984). Taking scale efficiency into account means that technical efficiency is estimated under the assumption that not all factories are operating at the optimal scale. The relationship between VRS and constant returns to scale (CRS) can be expressed as:

$$\textit{VRSTE Score}*\textit{Scale efficiency} = \textit{CRSTE Score} \quad \text{(A4)}$$

Taking further into consideration only the VRSTE score, the efficiency obtained to the scale is excluded. It is important for our analysis to leave out scale efficiency in order to provide a representative comparison of heterogeneous factories (see Kravtsova 2008 for a more detailed overview of the methodology and applications).

1.2. Physical distance

Bilateral great circle distances were calculated as follows. Let ϕ_j be the latitude and λ_j be the longitude information of factory j in radian form. Let j be inefficient factories and k the factories on the technological frontier with a maximum efficiency score, and let $\Delta\phi = \phi_k - \phi_j$ and $\Delta\lambda = \lambda_k - \lambda_j$ be the difference of the latitude and longitude, and R be the earth's radius (a radius of 6,371 kilometers was used). The great circle distance between j and k, d_{jk}, was calculated as follows:

$$d_{jk} = 2R \arcsin[\sin^2(\Delta\phi/2) + \cos\phi_j \cos\phi_k \sin^2(\Delta\lambda/2)]^{1/2}.$$

Appendix 2: Definition of Variables

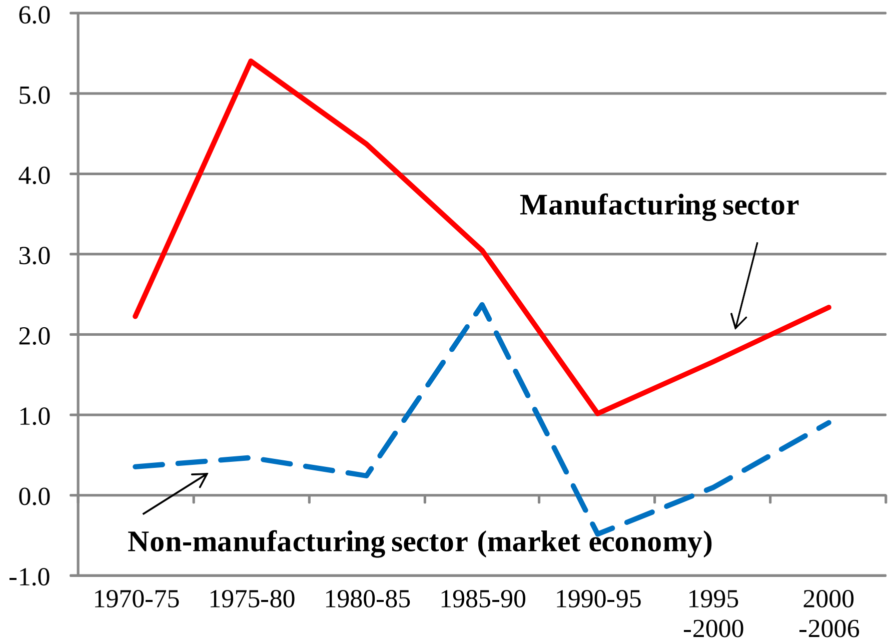
Variable	Description and Data Source	Formal Definition
Harris-distance from efficient (exporting) factories	Inverse distance weighted sum of efficient (exporting) factories (*Census of Manufactures*, 2007)	$Harris\_dis_j = \sum_{k \in E} \frac{1}{d_{j,k}}$ where j: non-efficient (non-exporting) factory; k: efficient (exporting) factory

Hanson-distance from efficient (exporting) factories	Inverse exponential distance weighted sum of efficient (exporting) factories (*Census of Manufactures*, 2007)	$Harris\_dis_j = \sum_{k \in E} \exp(-d_{j,k})$ where j: non-efficient (non-exporting) factory; k: efficient (exporting) factory
Factory size	Regular employment of the factory (*Census of Manufactures*, 2007)	Log of number of regular employees
Export dummy	Indicates whether a factory is involved in export activity or not (*Census of Manufactures*, 2007)	0/1 , where 1: exporter and 0: non-exporter
Multiple factory dummy	Dummy variable for factories that are part of multiple factories (*Census of Manufactures*, 2007)	0/1, where 1: multi-factory and 0: single-factory
Density	Population density of prefecture. Population divided by area of the prefecture (square kilometers), (*Population Census*, 2005)	Density_p=(Total population_p)/(area_p)
R&D expenditure	Prefectural government's R&D expenditure for institutions owned by local governments (million yen) over total population (*Population Census*, 2005)	$\text{R\&D}_p$=(R&D expenditure_p)/(Total population_p)
Share of university scholars	Share of university scholars in total number of workers, (*Population Census*, 2005)	University scholars_p=(Number of university scholars_p)/(Total number of workers)
Share of natural scientists	Share of natural scientists in total number of workers	Natural scientists_p=(Number of natural scientists_p)/(Total number of workers_p)

	(*Population Census*, 2005)	
Share of highly-educated people	Share of the number of university graduates in the total population of the prefecture (*Population Census*, 2005)	Highly educated$_p$=(Number of university graduates$_p$)/(Total population$_p$), where p: prefecture
Share of manager employment	Share of the number of managers in total employment in the prefecture (*Population Census*, 2005)	Managers$_p$=(Number of managers$_p$)/(Total employment$_p$) where p: prefecture
Share of technician employment	Share of the number of technicians in total employment in the prefecture (*Population Census*, 2005)	Technicians$_p$=(Number of technicians$_p$)/(Total employment$_p$) where p: prefecture
Share of university graduates' working hours	Share of total working hours by university graduate workers in working hours for all workers in the industry (*Population Census*, 2005)	University working$_i$=(University graduates' working time$_i$)/(Total working time$_i$), where i: industry
High-tech dummy	Takes value one if the industry is categorized as a high-tech industry according to OECD (2006)	0/1, where 1: high-tech industry and 0: other industry
Mid-high tech dummy	Takes value one if the industry is categorized as a mid-tech industry according to OECD (2006)	0/1, where 1: mid-high tech industry and 0: other industry
Low-tech dummy	Takes value one if the industry is categorizing as a low-tech industry according to OECD (2006)	0/1, where 1: low-tech industry and 0: other industry
Share of exports in prefecture's	Share of exports in the total shipments of the prefecture	Share of exports$_p$=(Total exports$_p$/Total shipments$_p$)

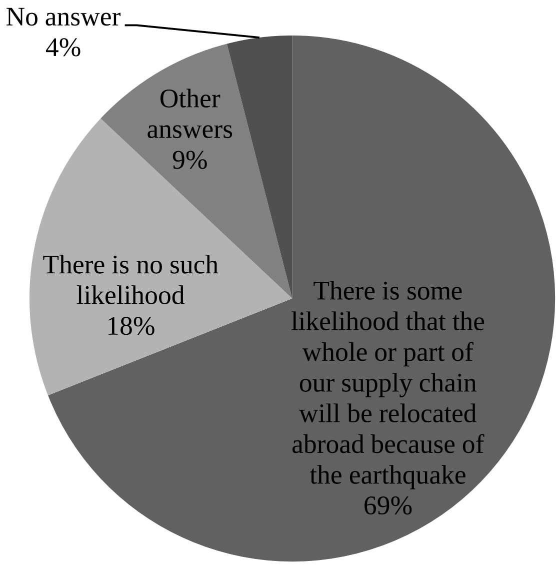
total shipments	(*Census of Manufactures*, 2007)	
Share of exports in industry's total shipments	Share of exports in the total shipments of the industry (*Census of Manufactures*, 2007)	Share of $exports_i$=(Total $exports_i$/Total $shipments_i$)
ln(Distance)* Exporter dummy	Log of Harris (Hanson) distance times exporter dummy	$ln(distance_j)\times$Exporter $dummy_j$

Figure 1. Japan's TFP Growth by Sector



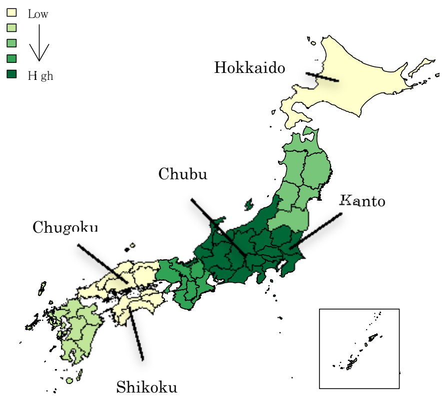
Source: Fukao and Kwon (2010).

Figure 2. Results of METI's Special Survey on Japan's Supply Chains after the Great East Japan Earthquake



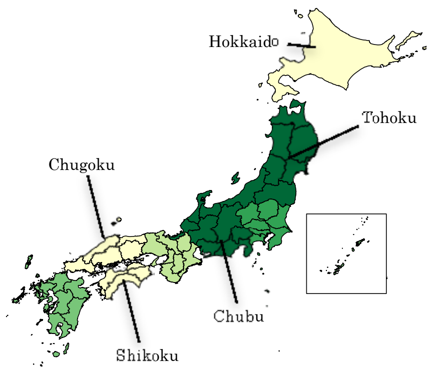
Source: "The Present Situation and Problems to Be Solved of the Japanese Economy after the Great East Japan Earthquake," METI, June 2011. The figure is based on METI's special survey on supply chains after the earthquake.

Figure 3. Distribution of Factories by Region

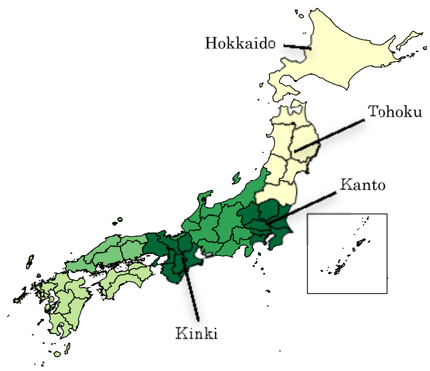


(a) Number of Factories

(b) Number of Efficient Factories



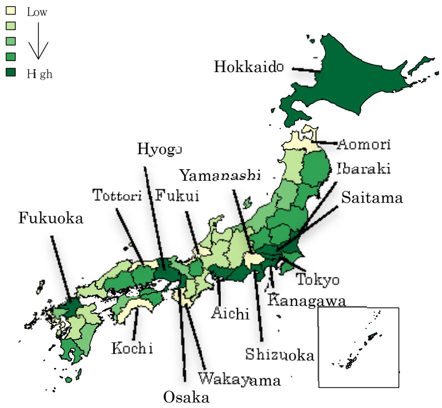
(c) No. of Efficient Factories /Total No. of Factories



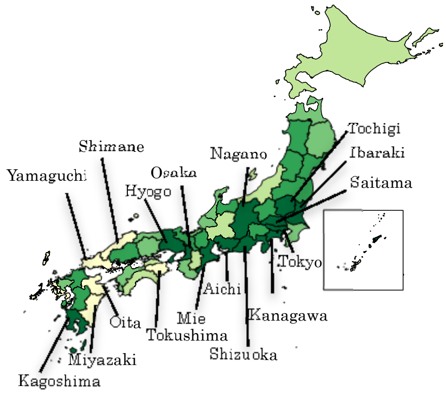
(d) No. of Efficient Factories /Area

Notes: A darker color implies a higher value for each measure. The corresponding values are shown in Table A1.

Figure 4. Distribution of Factories by Prefecture



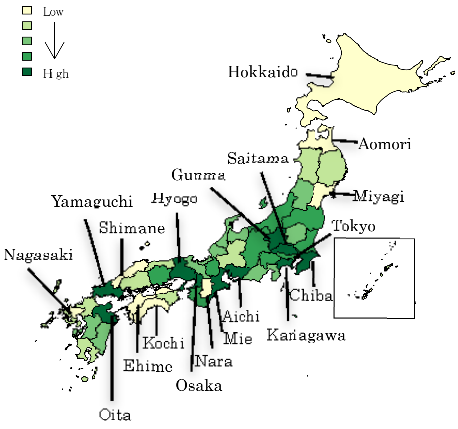
(a) Number of Factories



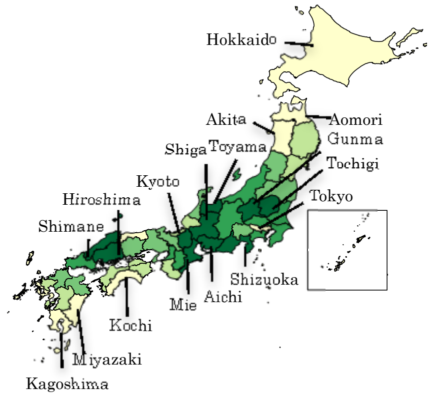
(b) Number of Efficient Factories

Note: A darker color implies a higher value for each measure. The corresponding values are shown in Table A2.

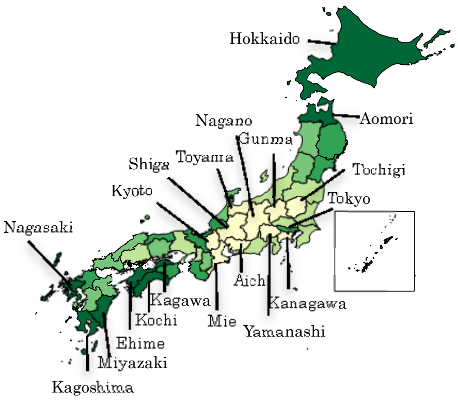
Figure 5. Distribution of Efficient Factories within Each of the Four Groups of Industries Based on Technological Intensity



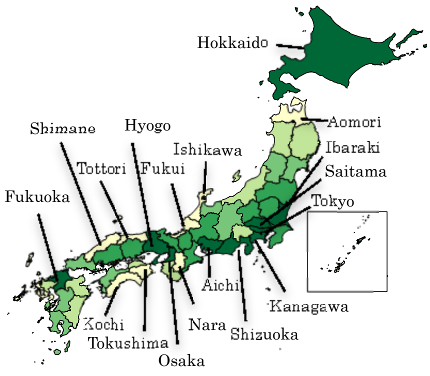
(a) High-Tech Factories



(b) Mid High-Tech Factories



(c) Mid Low-Tech Factories



(d) Low-Tech Factories

Note: A darker color implies a higher value for each measure. The corresponding values are shown in Table A3.

Figure 6. Kernel Density of Factories with Regard to Export Activity

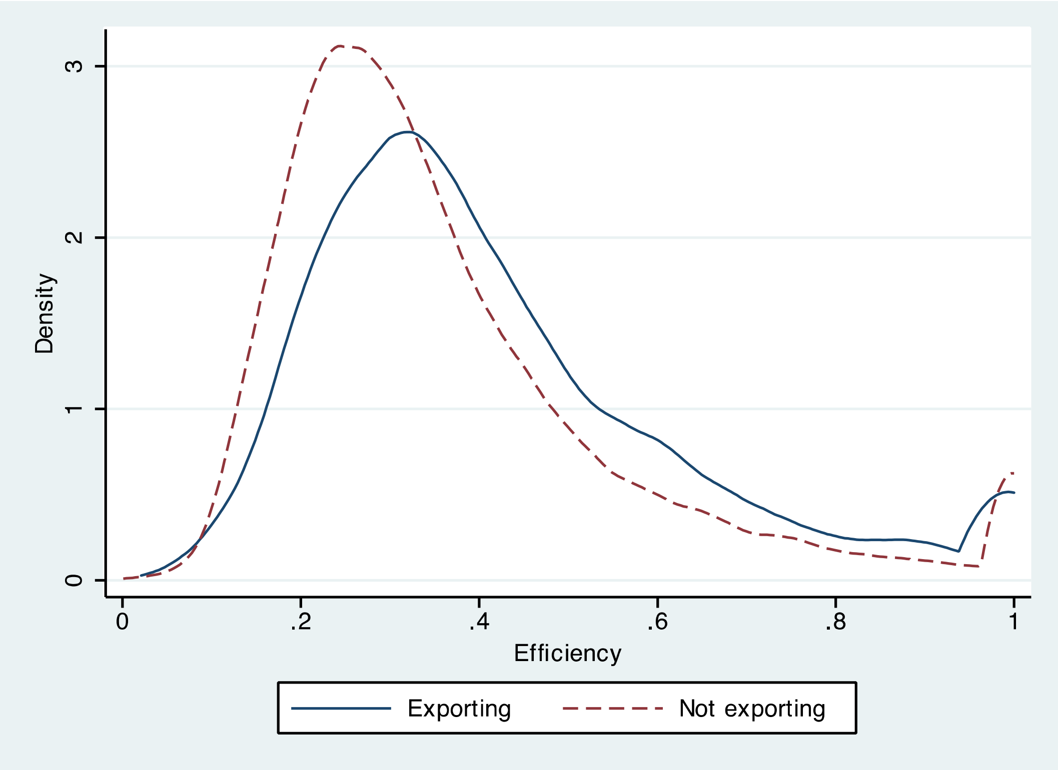
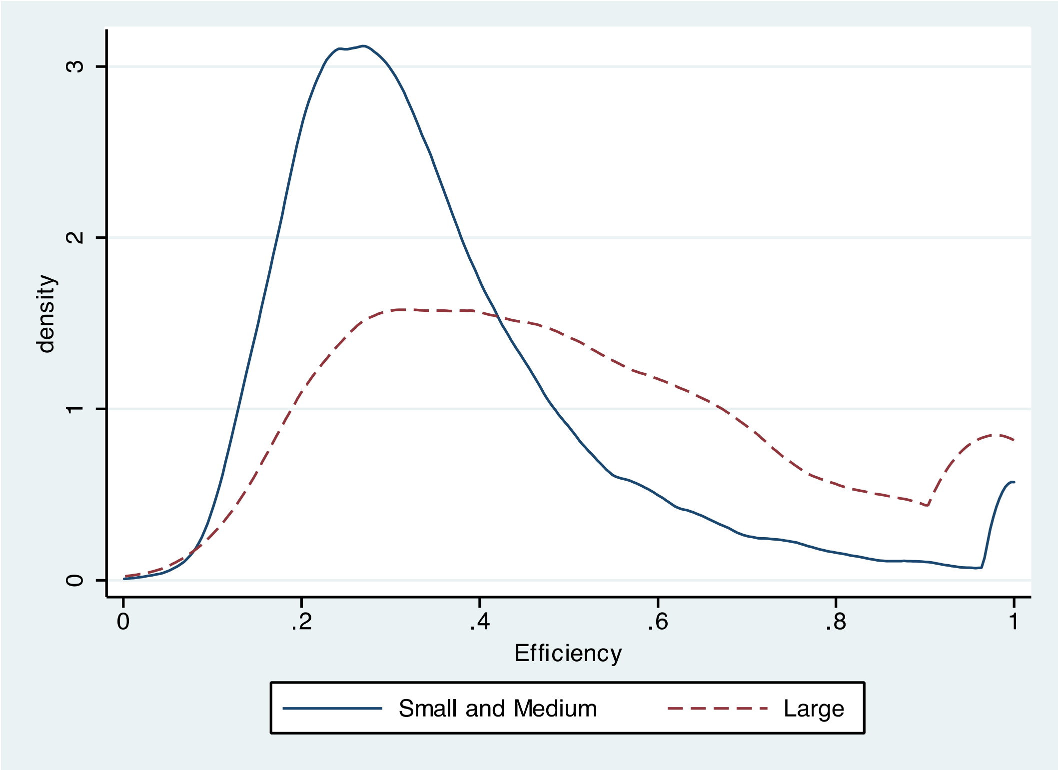


Figure 7. Kernel Density of Factories with Regard to Factory Size



Note: Following the definition of the Small and Medium-sized Enterprise Basic Act, small and medium factories are defined as factories whose regular workforce does not exceed 300 persons.

Figure 8. Kernel Density of Manufacturing Factories with Regard to Factory Number

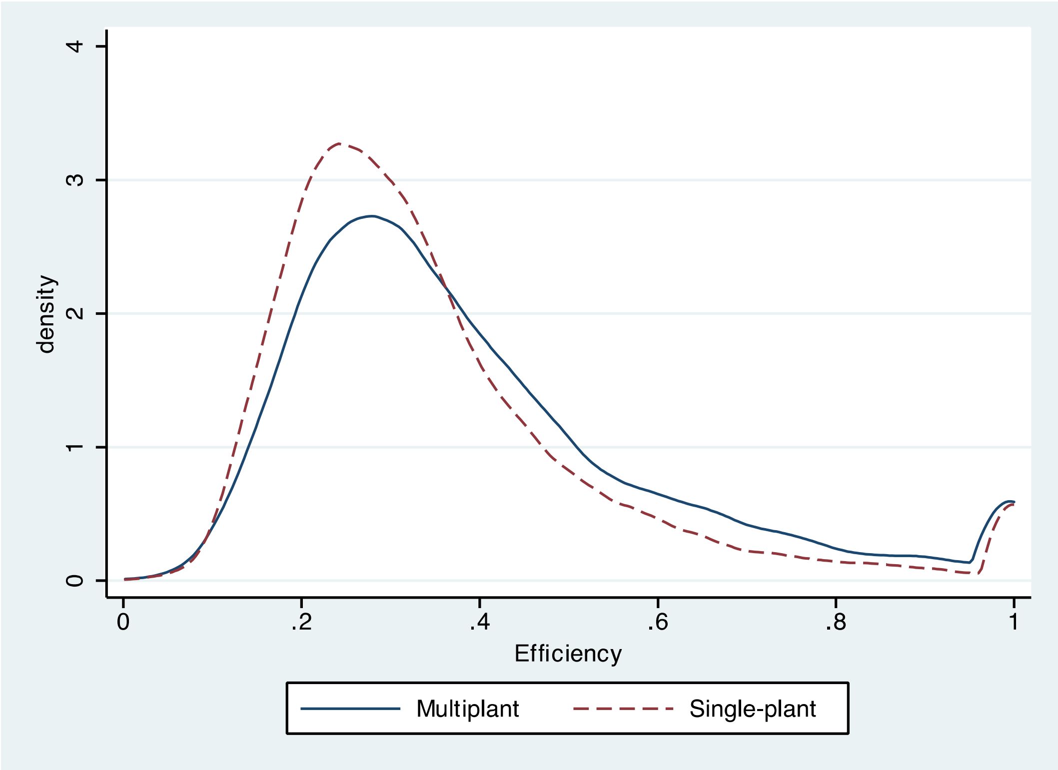


Table 1. Technical Efficiency and Distance to the Most Efficient Factories in the Prefecture: (Harris, 1954) and Hanson (2005) Distances Robustness Check

	Baseline_OLS	Baseline_truncated	Prefec_Industry_controls	Tech_Instensity_Dummy	Prefec_Industry_FE	EX_info	eBaseline_OLS	eBaseline_truncated	ePrefec_Industry_controls	eTech_Instensity_Dummy	ePrefec_Industry_FE	eEX_info
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ln(Harris-distance from efficient factories)	0.0102***	0.0124***	0.0046***	0.0042***	0.0037***	0.0036***						
	(12.0476)	*(12.3240)*	*(3.9576)*	*(3.8012)*	*(2.8722)*	*(3.2708)*						
ln(Hanson-distance from efficient factories)							0.0007***	0.0009***	0.0002*	0.0002*	0.0002***	0.0002**
							(8.4611)	*(7.3290)*	*(1.8062)*	*(1.7230)*	*(2.6435)*	*(2.0930)*
ln(Size)			0.0544***	0.0552***	0.0589***	0.0556***			0.0543***	0.0551***	0.0588***	0.0556***
			(44.2162)	*(43.4399)*	*(49.7628)*	*(44.8677)*			*(40.4112)*	*(43.2662)*	*(43.6620)*	*(41.9320)*
Exporter dummy			0.0135***	0.0150***	0.0209***	0.0192***			0.0133***	0.0147***	0.0210***	0.0191***
			(4.4602)	*(5.5111)*	*(7.0978)*	*(7.3534)*			*(4.4857)*	*(5.1495)*	*(7.3397)*	*(7.0500)*
Multi-plant_dummy			0.0218***	0.0218***	0.0212***	0.0204***			0.0219***	0.0218***	0.0212***	0.0205***
			(10.9843)	*(11.4080)*	*(11.6301)*	*(11.4496)*			*(12.1259)*	*(11.1877)*	*(11.8587)*	*(11.0885)*
ln(Density)			-0.0054**	-0.0053**		-0.0030			-0.0034	-0.0036		-0.0014
			(-2.2397)	*(-2.3370)*		*(-1.1832)*			*(-1.5351)*	*(-1.5817)*		*(-0.6185)*
ln(R&D expenditure)			-0.0065**	-0.0051		-0.0054			-0.0069**	-0.0055*		-0.0061*
			(-2.0129)	*(-1.4916)*		*(-1.4929)*			*(-1.9967)*	*(-1.6582)*		*(-1.7705)*
ln(Share of university scholars in the prefecture)			-0.0005	-0.0018		-0.0050			0.0009	-0.0006		-0.0046
			(-0.0969)	*(-0.4866)*		*(-1.3945)*			*(0.2181)*	*(-0.1398)*		*(-0.9942)*
ln(Share of natural scientists in the prefecture)			0.0023	0.0025		0.0025			0.0028	0.0029		0.0028
			(1.1261)	*(1.3403)*		*(1.2512)*			*(1.3170)*	*(1.4478)*		*(1.3746)*
ln(Share of highly-educated people)			0.0156*	0.0199**		0.0170*			0.0194*	0.0234**		0.0190**
			(1.6811)	*(2.0778)*		*(1.7869)*			*(1.9359)*	*(2.4188)*		*(2.0248)*
ln(Share of manager employment)			-0.0187	-0.0195		-0.0038			-0.0268**	-0.0267**		-0.0080
			(-1.3924)	*(-1.4409)*		*(-0.2665)*			*(-1.9774)*	*(-2.0216)*		*(-0.5908)*
ln(Share of technician employment)			0.0454**	0.0344*		0.0280			0.0352*	0.0252		0.0220
			(2.4476)	*(1.7616)*		*(1.5595)*			*(1.8998)*	*(1.2939)*		*(1.1550)*
ln(Share of univ. graduates' working hours)			-0.6498***	-0.9982***		-0.9289***			-0.6539***	-1.0022***		-0.9314***
			(-21.7509)	*(-20.6748)*		*(-20.4635)*			*(-21.6438)*	*(-20.4231)*		*(-20.3591)*
high_tech==2				0.0291***	0.0255***	0.0290***				0.0294***	0.0257***	0.0292***
				(6.8451)	*(5.8406)*	*(7.1415)*				*(6.7869)*	*(6.2762)*	*(6.7073)*
high_tech==3				0.0299***	0.0277***	0.0264***				0.0300***	0.0276***	0.0264***
				(8.1198)	*(6.9398)*	*(7.4496)*				*(7.7130)*	*(7.5308)*	*(7.3058)*
high_tech==4				0.0598***	0.0084**	0.0316***				0.0599***	0.0083**	0.0316***
				(12.8850)	*(2.0187)*	*(6.3459)*				*(11.9844)*	*(2.2302)*	*(6.6755)*
ln(Share of exporting in the prefecture)						-0.0076***						-0.0087***
						(-3.3560)						*(-4.0430)*
ln(Share of exporting in the industry)						-0.0118***						-0.0117***
						(-12.6685)						*(-14.0795)*
_cons	0.3472***	0.3321***	0.1612***	0.1246***	0.0896***	-0.1475***	0.3567***	0.3439***	0.1521***	0.1164***	0.0916***	-0.1692***
	(336.2581)	*(337.9862)*	*(3.4205)*	*(3.0187)*	*(10.5991)*	*(-2.9094)*	*(354.2836)*	*(326.2939)*	*(3.2193)*	*(2.6728)*	*(10.4969)*	*(-3.1989)*
/sigma		0.1839***	0.1709***	0.1704***	0.1715***	0.1696***		0.1842***	0.1710***	0.1704***	0.1715***	0.1697***
		(178.4808)	*(161.7225)*	*(166.9123)*	*(164.0700)*	*(169.5401)*		*(170.0096)*	*(162.6144)*	*(151.7528)*	*(171.2317)*	*(155.9144)*
Number of observations	42,042	42,042	39,387	39,383	40,713	39,358	41,968	41,968	39,314	39,310	40,640	39,285
R2	0.004						0.002					

Notes: *** p<0.01, ** p<0.05, * p<0.1.

t-statistics are provided in parentheses.

Dependent variable in all regressions: DEA score estimated with variable returns to scale model (VRSTE). Method of estimation: OLS in Model 1 and truncated regression in the rest of the models (2-10)

Model 1: Baseline linear regression. Model 2: baseline truncated regression. Model 3: controlling prefecture and industry-level variables.

Model 4: Model 3 + OECD technology intensity dummmies. Model 5: controlling prefecture fixed effects (FE). Model 6: includes share of exporting in the prefectures and in the industry.

Table 2. Distance to the Most Efficient Factories and Distance to Exporting Factories in Sectors with Different Technological Intensity

	High_Tech	High_Medium_Tech	Low_Medium_Tech	Low_Tech	exHigh_Tech	exHigh_Medium_Tech	exLow_Medium_Tech	exLow_Tech
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Harris-distance from efficient factories)	0.0004	0.0042**	0.0037*	0.0046				
	(0.0902)	*(2.1198)*	*(1.9119)*	*(1.6160)*				
ln(Size)	0.0372***	0.0540***	0.0619***	0.0671***	0.0298***	0.0448***	0.0610***	0.0653***
	(9.2077)	*(19.0923)*	*(31.7065)*	*(31.1390)*	*(6.5678)*	*(13.3817)*	*(27.3063)*	*(29.4851)*
Exporter dummy	0.0217**	0.0309***	0.0229***	0.0127*				
	(2.1326)	*(5.0751)*	*(5.2544)*	*(1.6730)*				
Multi-plant_dummy	0.0060	-0.0046	0.0124***	0.0467***	0.0082	-0.0016	0.0130***	0.0465***
	(0.7811)	*(-1.0338)*	*(4.1621)*	*(15.5329)*	*(1.1330)*	*(-0.3182)*	*(4.0880)*	*(14.8475)*
ln(Harris-distance from exporting factories)					-0.0048	0.0020	0.0025	0.0139***
					(-0.7199)	*(0.5300)*	*(0.6847)*	*(4.1770)*
_cons	0.1686***	0.0862***	0.1242***	0.0622***	0.1986***	0.1246***	0.1307***	0.0805***
	(3.8922)	*(3.7313)*	*(9.4675)*	*(5.8544)*	*(4.7717)*	*(5.0959)*	*(8.1573)*	*(7.6588)*
/sigma	0.1919***	0.1759***	0.1676***	0.1641***	0.1907***	0.1741***	0.1685***	0.1632***
	(54.8494)	*(73.5365)*	*(120.2333)*	*(90.5373)*	*(46.8672)*	*(65.6014)*	*(95.1940)*	*(89.8779)*
Number of observations	3,459	7,028	16,416	13,810	2,933	6,015	14,126	13,438

Notes: *** p<0.01, ** p<0.05, * p<0.1.

t-statistics are provided in parentheses.

Dependent variable in all regressions: DEA score estimated with variable returns to scale model (VRSTE).

Method of estimation: Truncated regression.

Models 1-4: restricting sample to high-tech (Model 1), mid-high tech (Model 2), mid-low tech (Model 3), and low-tech (Model 4) industry factories.

Models 5-8: distance to exporting establishments.

Table 3. Technical Efficiency and Distance to the Most Efficient Factories in Japan (Re-estimation of Table 1 for the Distances to the Most Efficient Factories in All Japan): Harris (1954) and Hanson (2005) Distances Robustness Check

	Baseline_OLS	Baseline_truncated	Prefec_Industry_controls	Tech_Instensity_Dummy	Prefec_Industry_FE	EX_info	eBaseline_OLS	eBaseline_truncated	ePrefec_Industry_controls	eTech_Instensity_Dummy	ePrefec_Industry_FE	eEX_info
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	coef/t
ln(Harris-distance from efficient factories)	0.0224***	0.0265***	0.0004	0.0018	0.0155***	0.0109***						
	(13.9818)	*(14.7781)*	*(0.1340)*	*(0.5370)*	*(3.6062)*	*(3.1396)*						
ln(Hanson-distance from efficient factories)							0.0010***	0.0012***	0.0003***	0.0003**	0.0004***	0.0003***
							(9.0578)	*(9.8603)*	*(2.6770)*	*(2.3594)*	*(3.1776)*	*(2.6772)*
ln(Size)			0.0544***	0.0552***	0.0589***	0.0557***			0.0544***	0.0552***	0.0588***	0.0556***
			(43.2064)	*(43.2713)*	*(44.5664)*	*(45.8797)*			*(41.9517)*	*(42.2729)*	*(49.6553)*	*(45.4619)*
Exporter dummy			0.0133***	0.0147***	0.0209***	0.0187***			0.0132***	0.0147***	0.0209***	0.0190***
			(4.6710)	*(5.4007)*	*(7.0051)*	*(6.5079)*			*(4.2995)*	*(5.2745)*	*(7.0288)*	*(6.6887)*
Multi-plant_dummy			0.0219***	0.0219***	0.0211***	0.0204***			0.0219***	0.0219***	0.0212***	0.0205***
			(11.8237)	*(10.9249)*	*(11.9992)*	*(11.0150)*			*(11.7146)*	*(10.8321)*	*(11.9604)*	*(10.3079)*
ln(Density)			-0.0033	-0.0038		-0.0039			-0.0037	-0.0038		-0.0017
			(-1.3016)	*(-1.5238)*		*(-1.5468)*			*(-1.5561)*	*(-1.5101)*		*(-0.7713)*
ln(R&D expenditure)			-0.0067**	-0.0053*		-0.0070**			-0.0070**	-0.0056*		-0.0062*
			(-2.0869)	*(-1.6719)*		*(-2.2128)*			*(-2.1948)*	*(-1.8264)*		*(-1.9486)*
ln(Share of university scholars in the prefecture)			0.0014	-0.0000		-0.0038			0.0008	-0.0007		-0.0047
			(0.3181)	*(-0.0054)*		*(-0.8534)*			*(0.2036)*	*(-0.1736)*		*(-1.1809)*
ln(Share of natural scientists in the prefecture)			0.0030	0.0029		0.0016			0.0027	0.0028		0.0027
			(1.5258)	*(1.3971)*		*(0.7981)*			*(1.4024)*	*(1.4699)*		*(1.4009)*
ln(Share of high-educated people)			0.0207**	0.0231**		0.0080			0.0192**	0.0232**		0.0188*
			(2.0926)	*(2.2624)*		*(0.7872)*			*(2.0075)*	*(2.4054)*		*(1.8161)*
ln(Share of manager employment)			-0.0266**	-0.0261**		0.0009			-0.0252*	-0.0253**		-0.0065
			(-2.1064)	*(-2.0019)*		*(0.0602)*			*(-1.9206)*	*(-1.9751)*		*(-0.5216)*
ln(Share of technician employment)			0.0313*	0.0242		0.0409**			0.0358*	0.0257		0.0227
			(1.6514)	*(1.3079)*		*(2.0355)*			*(1.8004)*	*(1.3150)*		*(1.1690)*
ln(Share of univ. graduates' working hours)			-0.6541***	-1.0039***		-0.9305***			-0.6516***	-1.0013***		-0.9304***
			(-22.6972)	*(-19.8595)*		*(-21.5046)*			*(-19.7477)*	*(-20.7316)*		*(-22.6671)*
high_tech==2				0.0294***	0.0254***	0.0290***				0.0293***	0.0255***	0.0290***
				(6.8472)	*(6.0271)*	*(6.5919)*				*(6.6371)*	*(6.3377)*	*(6.4331)*
high_tech==3				0.0301***	0.0276***	0.0263***				0.0300***	0.0276***	0.0264***
				(7.8910)	*(7.1304)*	*(6.6099)*				*(7.2403)*	*(7.2090)*	*(7.0745)*
high_tech==4				0.0602***	0.0084**	0.0316***				0.0600***	0.0085**	0.0316***
				(12.5944)	*(2.1922)*	*(6.1897)*				*(12.3346)*	*(2.1858)*	*(6.6629)*
ln(Share of exporting in the prefecture)						-0.0109***						-0.0085***
						(-4.3786)						*(-3.6036)*
ln(Share of exporting in the industry)						-0.0118***						-0.0118***
						(-15.0250)						*(-14.2602)*
_cons	0.3076***	0.2854***	0.1496***	0.1173**	0.0759***	-0.1784***	0.3579***	0.3454***	0.1534***	0.1176***	0.0936***	-0.1661***
	(91.3288)	*(75.2414)*	*(3.2005)*	*(2.5683)*	*(8.2294)*	*(-3.2964)*	*(361.1139)*	*(342.9054)*	*(3.6643)*	*(2.7046)*	*(10.9229)*	*(-3.3428)*
/sigma		0.1838***	0.1709***	0.1704***	0.1714***	0.1696***		0.1841***	0.1709***	0.1704***	0.1715***	0.1696***
		(180.0196)	*(155.2541)*	*(180.8200)*	*(167.4558)*	*(165.1070)*		*(178.6539)*	*(165.0349)*	*(186.1318)*	*(168.0130)*	*(169.8362)*
Number of observations	42,042	42,042	39,387	39,383	40,713	39,358	42,042	42,042	39,387	39,383	40,713	39,358
R2	0.004						0.002					

Notes: *** p<0.01, ** p<0.05, * p<0.1.

t-statistics are provided in parentheses.

Dependent variable in all regressions: DEA score estimated with variable returns to scale model (VRSTE). Method of estimation: OLS in Model 1 and truncated regression in the rest of the models (2-10)

Model 1: baseline linear regression. Model 2: baseline truncated regression. Model 3: controlling for prefecture and industry-level variables.

Model 4: Model 3 + OECD technology intensity dummmies. Model 5: controlling for prefecture fixed effects. Model 6: includes the share of exporting in the prefecture and in the industry.

Table 4. Distance to the Most Efficient Factories and Distance to Exporting Factories in Sectors with Different Technological Intensity (Re-estimation of Table 2 for Distances within the Whole Country)

	High_Tech	High_Medium_Tech	Low_Medium_Tech	Low_Tech	exHigh_Tech	exHigh_Medium_Tech	exLow_Medium_Tech	exLow_Tech
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Harris-distance from efficient factories)	-0.0091	0.0100	0.0152**	0.0273***				
	(-0.6148)	*(1.2699)*	*(2.3996)*	*(2.8159)*				
ln(Size)	0.0373***	0.0540***	0.0620***	0.0671***	0.0297***	0.0449***	0.0611***	0.0654***
	(9.3833)	*(19.6139)*	*(35.9494)*	*(29.2305)*	*(6.7190)*	*(14.8993)*	*(26.2647)*	*(27.8635)*
Exporter dummy	0.0217**	0.0310***	0.0229***	0.0127*				
	(2.0754)	*(4.9954)*	*(5.6960)*	*(1.7489)*				
Multi-plant_dummy	0.0060	-0.0046	0.0124***	0.0465***	0.0081	-0.0016	0.0131***	0.0463***
	(0.7541)	*(-1.0919)*	*(4.3597)*	*(14.2253)*	*(0.9989)*	*(-0.3293)*	*(4.3564)*	*(14.3072)*
ln(Harris-distance from exporting establishments)					-0.0348	-0.0000	0.0158**	0.0354***
					(-1.4868)	*(-0.0022)*	*(1.9764)*	*(3.6554)*
_cons	0.1756***	0.0767***	0.1105***	0.0394***	0.2626***	0.1232***	0.1017***	0.0057
	(4.0446)	*(3.1609)*	*(6.7860)*	*(3.0756)*	*(4.3370)*	*(3.9819)*	*(5.1229)*	*(0.2978)*
/sigma	0.1919***	0.1760***	0.1676***	0.1641***	0.1906***	0.1741***	0.1685***	0.1632***
	(49.5529)	*(72.1438)*	*(108.4984)*	*(84.3325)*	*(48.8826)*	*(59.9462)*	*(96.9286)*	*(85.6842)*
Number of observations	3,459	7,028	16,416	13,810	2,933	6,015	14,126	13,438

Notes: *** p<0.01, ** p<0.05, * p<0.1.

t-statistics are provided in parenthes.

Dependent variable in all regressions: DEA score estimated with variable returns to scale model (VRSTE).

Method of estimation:truncated regression.

Models 1-4: restricting sample to high-tech (Model 1), mid-high tech (Model 2), mid-low tech (Model 3), and low-tech (Model 4) industry factories.

Models 5-8: distance to exporting establishments.

Table A1. Data Underlying Figure 3

	Figure 3(a)	Figure 3(b)	Figure 3(c)	Figure 3(d)
Region	No. of factories	No. of most efficient factories	No. of most efficient factories/ Total number of factories	No. of most efficient factories/ Area
Hokkaido	1,157	44	0.038	0.053
Tohoku	4,151	121	0.029	0.181
Kanto	10,762	328	0.030	1.012
Chubu	11,084	283	0.026	0.424
Kinki	7,775	261	0.034	0.788
Chugoku	3,087	96	0.031	0.301
Shikoku	1,402	52	0.037	0.276
Kyushu	3,955	146	0.037	0.328

Table A2. Data Underlying Figure 4

	Figure 4(a)	Figure 4(b)
	No. of factories	No. of most efficient factories
Hokkaido	1,157	44
Aomori	387	12
Iwate	676	14
Miyagi	639	26
Akita	519	16
Yamagata	753	22
Fukushima	1,177	31
Ibaraki	1,475	59
Tochigi	1,103	29
Gumma	1,157	31
Saitama	2,265	60
Chiba	1,194	34
Tokyo	1,633	44
Kanagawa	1,935	71
Niigata	1,057	16
Toyama	763	24
Ishikawa	601	11
Fukui	432	8
Yamanashi	468	15
Nagano	1,315	20
Gifu	1,247	28
Shizuoka	1,619	53
Aichi	3,582	108
Mie	967	39
Shiga	840	36
Kyoto	826	27
Osaka	2,594	62
Hyogo	1,844	72
Nara	387	11
Wakayama	317	14
Tottori	264	5

Shimane	262	3
Okayama	916	27
Hiroshima	1,090	39
Yamaguchi	555	22
Tokushima	258	9
Kagawa	446	14
Ehime	512	23
Kochi	186	6
Fukuoka	1,341	53
Saga	378	13
Nagasaki	314	8
Kumamoto	531	19
Oita	386	15
Miyazaki	389	14
Kagoshima	451	18
Okinawa	165	6

Table A3. Data Underlying Figure 5

	Figure 5(a)	Figure 5(b)	Figure 5(c)	Figure 5(d)
Prefecture	No. of factories (High-tech industry)	No. of factories (Mid high-tech industry)	No. of factories (Mid low-tech industry)	No. of factories (Low-tech industry)
Hokkaido	56	83	260	758
Aomori	57	38	84	208
Iwate	85	87	227	277
Miyagi	84	77	204	274
Akita	101	66	159	193
Yamagata	101	114	271	267
Fukushima	177	218	432	349
Ibaraki	118	261	644	449
Tochigi	112	187	486	318
Gunma	96	236	535	290
Saitama	165	435	932	733
Chiba	60	256	488	390
Tokyo	165	296	496	676
Kanagawa	175	462	840	457
Niigata	106	186	431	334
Toyama	101	115	349	198
Ishikawa	49	91	225	236
Fukui	38	66	140	188
Yamanashi	78	82	180	128
Nagano	236	232	542	305
Gifu	89	185	649	324
Shizuoka	117	259	723	520
Aichi	154	678	1,857	893
Mie	83	193	479	212
Shiga	95	148	412	185
Kyoto	57	125	314	330
Osaka	167	603	1,032	792
Hyogo	144	421	721	558
Nara	30	47	156	154
Wakayama	17	56	109	135
Tottori	54	34	62	114
Shimane	25	22	116	99

Okayama	70	162	355	329
Hiroshima	71	163	496	360
Yamaguchi	25	107	234	188
Tokushima	18	38	82	120
Kagawa	29	50	145	222
Ehime	20	60	173	259
Kochi	16	20	54	96
Fukuoka	62	187	531	561
Saga	35	36	126	181
Nagasaki	24	40	87	163
Kumamoto	44	92	191	203
Oita	24	82	148	132
Miyazaki	27	60	121	181
Kagoshima	49	58	75	269
Okinawa	3	18	37	107