



RIETI Discussion Paper Series 10-E-017

# **Productivity Distribution, Firm Heterogeneity, and Agglomeration: Evidence from firm-level data**

**OKUBO Toshihiro**

Kobe University

**TOMIURA Eiichi**

RIETI



Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry

<http://www.rieti.go.jp/en/>

**Productivity Distribution, Firm Heterogeneity, and Agglomeration:  
Evidence from firm-level data**

Toshihiro OKUBO

Kobe University

and

Eiichi TOMIURA

Research Institute of Economy, Trade and Industry, and Yokohama National University

Abstract

This paper empirically examines how productivity distributions of firms vary across regions based on Japan's manufacturing census data. We confirm the established finding of higher average productivity in core regions, but find that firm productivity is distributed with wide dispersions, especially in core regions. Our firm-level estimates demonstrate that the productivity distribution of firms tends to be noticeably left-skewed deviating from the normal distribution, especially in regions with weak market potential but also in agglomerated or urbanized regions. These findings suggest that agglomeration economies are likely to accommodate heterogeneous firms to co-exist in the same region.

Key words: agglomeration; productivity; gamma distribution; heterogeneity; firm-level data

JEL classification: L11; R12

RIETI Discussion Papers Series aims at widely disseminating research results in the form of professional papers, thereby stimulating lively discussion. The views expressed in the papers are solely those of the author(s), and do not present those of the Research Institute of Economy, Trade and Industry.

## 1. Introduction

Productivity differs across firms subject to a certain distribution. At an aggregated regional level, average productivity in big cities and agglomerated areas tends to be substantially high.<sup>1</sup> These facts have been separately investigated, but little is known about how the firm productivity distribution is affected by geographical factors. This paper empirically examines, based on firm-level manufacturing census data, how higher moments (skewness and dispersion) of the productivity distribution across firms vary depending on the level of agglomeration.

Firms are tremendously heterogeneous in productivity even within the same region.<sup>2</sup> In a notable study on the firm productivity dispersion, Syverson (2004a) argues that larger local demand leads to a productivity distribution truncated from below due to intensified competition and finds empirical evidence consistent with this prediction in the case of the ready-made concrete industry. However, factors other than intense competition are likely to affect the shape of productivity distributions. Among them the agglomeration effect and Marshallian externality should be critical in considering economic geography. If these effects dominate the competition effect firms should distribute over wider ranges of productivity in agglomerated regions, by accommodating unproductive firms thus allowing them to survive in the regions. Agglomeration fosters more varieties of products as well as wider ranges of firm productivity. In particular, many small-sized suppliers with relatively low productivity may operate in close proximity to a large, productive final assembler by providing inputs tailored to complicated assembler's requirements, probably facilitated by face-to-face contacts and local knowledge spillovers. Larger local demand may also allow heterogeneous firms to survive in the same region by supporting wider varieties of product differentiation. In other words, agglomeration should

---

<sup>1</sup> See Rosenthal and Strange (2004) and Melo et al. (2009) as useful surveys of previous work.

<sup>2</sup> This paper focuses on the firm side, but heterogeneity is also an important issue on the worker side. Using French data, Combe et al. (2008) investigate spatial selection in heterogeneous workers.

allow wider ranges of heterogeneous firms through positive externalities experienced in the world of many differentiated products as in advanced economies of our age.

In a different context, Cabral and Mata (2003), using Portuguese manufacturing census data, report that the firm size distribution is substantially right-skewed and becomes more proximate to a log-normal distribution as firms get older. Our paper is a spatial parallel to Cabral and Mata (2003) in that both estimate higher moments of the distribution of firms to examine its relation to competition. As empirical investigations of productivity in the context of agglomeration have been seriously limited, with Syverson (2004a) as a rare example, this paper should represent an important input.

This paper investigates the productivity premium of agglomeration, as discussed in the economic geography and urban economics literature, as well as the productivity distribution of firms across regions as an application of the industrial organization literature, such as Cabral and Mata (2003). We empirically investigate the shape of productivity distributions based on firm-level data derived from Japan's manufacturing census. All firms with no less than five employees in all manufacturing industries across all regions in Japan are included in our sample of six consecutive waves of censuses. To preview the principal results, this paper first confirms the established finding of an agglomeration effect on productivity: the average productivity is sizably higher in core regions. Second, the productivity of firms tends to be distributed over relatively wide ranges, obviously deviating from the normal distribution, especially in core regions. Third, by linking the estimated parameters of a gamma distribution with economic geography variables, we find that the productivity distribution tends to be less left-skewed (closer to the normal distribution) in regions with stronger market potential. The deviation from the normal distribution is sustained in agglomerated or urbanized regions, suggesting the important role of the positive externality in shaping the distribution of productivity across firms.

The rest of this paper is organized as follows. Section 2 describes the data. Section 3 reports our empirical results on the distributions of firm-level productivity and relates the parameter estimates to economic geography. Section 4 discusses the implications of our results. Section 5 adds concluding comments.

## **2. Data description**

This section is devoted to the explanations of our micro-data derived from Japan's *Census of Manufacturers*. This census covers virtually all plants across all manufacturing industries.<sup>3</sup>

Although the annual survey covers plants above the given size threshold, small-sized plants are included only in the "census years" (years with a 0, 3, 5, or 8 as its last digit). As the principal purpose of this paper is the investigation of productivity distributions over the entire population of plants, we concentrate on census years to avoid truncations due to the sampling of plants. While plants of any size, including those with only one employee, are covered by the census, plant-level data are maintained only for the plants with no less than five employees in the original micro-data files of the central government even for the most recent census. As a result, our sample excludes plants with less than five employees. Since these extremely small-sized plants produce negligible volumes of output, their omission is unlikely to affect our conclusion on economic geography.

Our sample consists of the following six census years: 1978, 1980, 1983, 1985, 1988, and 1990, since the plant-level data before the mid-1970s are no longer available, even from the original government data files. By using these six consecutive waves of manufacturing censuses, we can investigate the productivity distributions over Japan's history from the oil crises (1973) to the bubble economy (several years prior to 1992). We decide to focus on these earlier years

---

<sup>3</sup> Henderson (2003) studied Marshallian externality based on U.S. *Census of Manufacturers*.

from the following reasons. From the 1990s onward, plant location decisions by Japanese firms have become increasingly global due to expanded production overseas (in particular toward Asia) by Japanese multinationals, accelerated by the unprecedented exchange rate appreciation. No information on offshore production is available in the domestic manufacturing census. By contrast, the 1970s and 1980s, which are the focus of our paper, experienced a transition from high-speed to steady growth. Importantly, the Japanese economy in this period did not experience substantial foreign direct investment, offshoring or international outsourcings. In parallel, this period corresponds to the transition from the bi-polar urban system driven by Tokyo and Osaka to the mono-polar urban system leading to mega-concentration in Metropolitan Tokyo, as suggested by Fujita and Tabuchi (1997). Therefore, the period of the 1970s and 80s, which is our data sample, involves many interesting questions on spatial patterns of firm location and is an appropriate period over which to investigate relationships among firm location, firm productivity and market competition without taking into account overseas production and hollowing-out. As no plant identifier tracing micro-data over time is available, our data set is unfortunately in the format of repeated cross-sections. Since the main target of this paper is the comparison of productivity between core and periphery regions, not on the entry-exit dynamics of plants, this data limitation is unlikely to affect our principal conclusions.

The manufacturing census contains basic information on plant-characteristics, such as output (shipment) and employment (number of regular workers). Whether or not each plant is a part of a multi-plant firm is also reported, though no identifier is available for linking plants under the same ownership. Hence, the aggregation of our plant-level data to the firm level is impossible from our census data. Since a plant location decision should be affected by the locations of other plants owned by the same firm in the case of multiple-plant firms, we concentrate on the sample of single-plant firms for investigating the distribution of productivity.

Since single-plant firms occupy the substantial share in the population of plants (74.5% in 1990), the exclusion of multi-plant firms does not affect our principal results. Our sample of single-plant firms contains as many as 324,687 firms in 1990. By concentrating on single-plant firms, we use “firm” and “plant” interchangeably below. Appendix Table A presents basic summary statistics of our census data.

### **3. Empirical results**

#### **3.1. Comparisons of average productivity across regions**

Before investigating the productivity distributions, this section examines how the average productivity of firms located in agglomerated core regions differ from that in peripheral regions.

The territory of Japan is divided into 47 prefectures, each of which roughly corresponds to a NUTS2 region.<sup>4</sup> To identify the agglomeration effect, we focus on the three prefectures with the biggest population: Tokyo, Osaka, and Aichi. These prefectures are obvious economic centers and the core regions of Japan, as they account for around 32 percent of industrial output, 26 percent of manufacturing output, 32 percent of GDP, and 22 percent of the population of Japan in 2005. To check the robustness of our focus on these three prefectures, we also examine the Greater Tokyo Area and the Greater Osaka Area by including neighboring prefectures<sup>5</sup>. This paper defines these regions (Tokyo, Osaka, Aichi, Greater Tokyo, and Greater Osaka) as the core region (Core) and the others as the periphery region.

Figures 1-a to 1-g report histograms of productivity distribution (frequency is in terms of the logarithm of firm productivity) for all firms in Japan combined in (1-a), firms in Tokyo (1-b),

---

<sup>4</sup> See Appendix Table B for prefecture name and code.

<sup>5</sup> We define Greater Tokyo Area (nation capital area, or *shuto-ken* in Japanese) as Tokyo and neighboring prefectures: Kanagawa, Chiba and Saitama. Greater Osaka Area (Kyoto-Osaka-Kobe Area, or *Keihanshin* in Japanese) is defined as Osaka and the neighboring Kyoto and Hyogo prefectures.

Greater Tokyo Area (1-c), Aichi (1-d), Osaka (1-e), Greater Osaka Area (1-f) and Core (1-g), respectively. Productivity is measured by per-worker value-added, since it is practically impossible to estimate the total factor productivity of each firm in our repeated cross-section data set without any longitudinal identifier. We note that as some previous studies discussed, the productivity measured by per-worker value-added are not crucially different from other productivity measures (e.g. TFP).<sup>6</sup> The average productivity in all of the core prefectures is clearly higher than that of total Japanese firms.

We also test whether the average of the productivity distribution is significantly different between core and peripheral regions, based on the Kolmogorov-Smirnov (KS) test, a non-parametric technique. The KS test first distinguishes two groups of core and periphery and then measures the difference in averages between two groups. As Table 1 reports, firm productivity in Tokyo (Greater Tokyo Area and Greater Osaka Area) is 18%-19% (15-16%, 6-9%, respectively) higher than in peripheral prefectures. More generally, the productivity in the core region, which is defined by a composite of the Greater Tokyo, Greater Osaka and Aichi areas, is 16 to 17 % higher than that in the other prefectures. All results are statistically significant with zero associated p-values.

Next, we investigate the impact of agglomeration economies on productivity in depth, taking into account region-specific factors and/or sector-specific factors. To evaluate the magnitude of agglomeration economies in firm productivity, we estimate the following regression:

$$PROD_{ij} = c + \sum \alpha_k CORE_k + \sum \beta_l SECTOR_l + \sum \gamma_m PREF_m + \varepsilon_i \quad (1)$$

where  $PROD_{ij}$  denotes the logarithm of firm productivity for firm  $i$ , located in prefecture  $j$ .

---

<sup>6</sup> The results are unlikely to be qualitatively affected by the choice of productivity measures. See Bernard and Jones (1996), for example.



$CORE \in \{(Tokyo, Osaka, Aichi), (Greater Tokyo, Greater Osaka), (all\ core\ regions\ combined)\}$  is a vector of core region dummy variables, where if firm  $i$  is located in Tokyo, the *Tokyo* dummy is one and zero otherwise. The sector dummy *SECTOR* is defined for 2-digit level manufacturing. *PREF* is a prefecture dummy. If firm  $i$  is located in prefecture  $j$  (i.e.  $m=j$ ), then the prefecture  $j$  dummy is one and otherwise zero. The error term and constant terms are expressed by  $\varepsilon$  and  $c$  respectively in (1).

The OLS regression results from (1) are summarized in Table 2. In all cases reported in this table, the dummy variables for the core regions are statistically significant at any conventional significance level. The productivity of average plants located in these core regions tends to be remarkably higher than that in peripheral regions. The productivity premium due to agglomeration is sizable: between 20% and 50%.

The magnitude of the agglomeration effect on productivity reported here is compared with previous results, such as 4.5% in Europe at NUTS3 region level reported by Ciccone (2002), 13% at NUTS2 level in Europe from Brühlhart and Mathys (2008), 3-8% surveyed by Rosenthal and Strange (2004), and 5.8%, estimated based on meta-analysis by Melo et al. (2009). Our finding is also higher than previous results for other Japanese data, such as Dekle (2002) and Nakamura (2008), which found productivity was around 15% higher in agglomerated regions. As Strange (2009) pointed out in his survey of agglomeration estimates, one of the possible reasons for this high estimate is likely to be at least partly due to cross-regional variations in human capital, which we cannot control for within our micro data.<sup>7</sup> In what follows, this paper focuses on the shape (higher moments) of the productivity distribution, rather than exploring underlying causes of the average productivity premium.

---

<sup>7</sup> Japanese industrial statistics in most cases, including our census data, do not contain employment data disaggregated by skills, occupations, or educational attainment.

### 3.2. Distributions of firm productivity

While the previous section compared mean productivity at the regional level averaged over heterogeneous firms, we cannot ignore the shape of the productivity distribution, i.e. dispersions and skewness across firms located in the same region. The productivity distributions are analyzed first by visual inspections of distribution graphs and then by estimations the parameters of gamma distributions.

A brief consideration of the productivity distributions displayed in Figure 1 is informative. As the frequency of firms within each productivity interval is measured on the vertical axis, each histogram can be regarded as an empirical counterpart of the probability density. Figure 1-a covers all regions in Japan, while Figure 1-b, 1-d, and 1-e present the corresponding distributions for firms located in Tokyo, Aichi, and Osaka, respectively. Visual inspection of these histograms indicates that firms located in the three core regions tend to distribute over a wider range of productivity compared with the national average. This finding appears to be in contrast to the greater productivity dispersion in smaller local markets observed by Syverson (2004a) in the case of U.S. ready-made concrete.<sup>8</sup> While Syverson (2004a) argues that intensified competition through cross-product substitution in larger local markets truncates the productivity distribution from below, this paper will examine whether or not other factors, such as externalities, are related to the shape of productivity distribution in the next section.

From the density histograms we also note that the distributions appear to obviously deviate from the normal distribution and are left-skewed. To check the validity of such an impression, we first calculate Kernel density estimates, Figure 2-a to 2-c present the results of this approach. We find that the productivity density is not distributed (log-) normal, but is

---

<sup>8</sup> Syverson (2004b) compares 443 U.S. manufacturing industries to complement Syverson (2004a), and finds less productivity dispersion in industries with high substitutability, which is proxied by a value/weight ratio or shipped distance.

definitely left-skewed (Figure 2-a). The distribution in core regions appears more left-skewed than that in all of Japan (Figure 2-b). It is also clear that the distribution in core regions is more skewed than that of peripheral regions (Figure 2-c). We conclude that the distribution of firm productivity in peripheral regions is relatively close to the log-normal distribution while that in core regions is more left-skewed.

Although the Kernel density graphs in Figure 2 clearly demonstrate the core-periphery differences, we cannot exclude the possibility that these cross-regional variations in productivity distributions may be merely due to differences in industrial compositions across regions (due to higher share of high-productivity industries being located in core regions).<sup>9</sup> To check this possibility, we present Kernel density estimates for major industries in Figure 3. These industry-specific results confirm that our previous finding is not entirely explained by differences in industrial compositions. Productivity remains on average higher and its distributed is still more left-skewed in core regions than that in periphery even within each industry, though we cannot neglect industrial structure altogether as the core-periphery difference naturally varies in magnitude depending on the industry.<sup>10</sup>

Next, we test whether the productivity distribution is (log-) normal or not, using skewness and kurtosis statistics. As a result, the log normality tests for productivity in all prefectures are significantly rejected. Thus we can confirm that the distribution of productivity is *not* log-normal.

Many empirical studies on the distribution of firm size have shown that firm size is subject to log-normal distribution following Gibrat's law. However, recent studies using plant-level data sets, including small business, have derived different outcomes. For example,

---

<sup>9</sup> Holmes and Stevens (2002) show the strong connection between firm size and industry concentration.

<sup>10</sup> In analyzing how much the average productivity differs across regions, we should not ignore cross-regional variations in industrial structures. See the regression results in Appendix Table D.

Cabral and Mata (2003) find that firm distribution is not log-normal and is skewed toward smaller sizes (“right-skewed”), however it evolves over time toward log-normal distribution as firms age. Our finding shows that firm productivity is again not distributed log-normally and is left-skewed, but becomes more left-skewed as regions are more agglomerated. Before comparing the differences we note that the firm distribution in Cabral and Mata (2003) and other empirical studies is measured in firm size (e.g. employee and profit) rather than in productivity.<sup>11</sup>

As investigated by Cabral and Mata (2003), this paper estimates the extended generalized gamma distribution with a probability density function defined as follows:

$$\frac{|\kappa|}{\Gamma(\kappa^{-2})} (\kappa^{-2})^{-2} \exp(\kappa^{-2}(\kappa q - \exp(\kappa q))) \quad \kappa \neq 0 \quad (2)$$

$$\frac{1}{\sqrt{2\pi}} \exp(-1/2q^2) \quad \kappa = 0 \quad (3)$$

where  $q \equiv (\ln prod - \mu)/\sigma$  is a function of firm productivity,  $prod$ ,  $\mu$  is its mean,  $\sigma$  its standard deviation and  $\kappa$  is the shape parameter of the gamma distribution.  $\Gamma$  denotes the gamma function. As shown in Figure 4, when  $\kappa$  goes to zero the distribution is (log-) normal distribution, as specified in (3). When  $\kappa$  is more (less) than zero in (2), the distribution is left-skewed (right-skewed).

We now estimate the firm productivity distribution in each prefecture for each year. Table 3 and Figure 5 report the estimation results of  $\kappa$  and  $\sigma$ . All of the  $\kappa$ 's are significantly positive (varying in value from 0.3 to 0.8), while  $\sigma$  takes a value around three. This tells us that firm productivity distributions are left-skewed in all 47 prefectures. Given  $\sigma$ , larger positive value of

---

<sup>11</sup> Barrios, et al. (2005), using Irish manufacturing census data, discovered a firm distribution skewed by financial constraints, but Angelini and Generale (2005), using Italian survey data, suggest no impact of financial constraints on firm distribution. More generally, Angelini and Generale (2008) found that financial constraints have no significant impact on the evolution of firm distribution in OECD countries.

$\kappa$  (more left-skewed) means that firms with relatively low productivity are more likely to survive, while smaller  $\kappa$  (close to normal distribution) indicates that less productive firms are pushed out possibly due to severe local competition.

We find several interesting outcomes from comparisons across regions.<sup>12</sup> First, the shape parameter  $\kappa$  is quite heterogeneous across regions (the upper panel of Figure 5). While periphery regions geographically far from core regions often exhibit high values (e.g. 0.6-0.7 in Hokkaido, Aomori and Oita), the values in Tokyo and other core regions are not the smallest observed values (the value is around 0.55 in Tokyo). When considering  $\sigma$  we note that the cross-regional variations are much smaller than in  $\kappa$  (see the middle panel of Figure 5), but core regions tend to have slightly higher values of  $\sigma$ . These results, which are richer than those in Syverson (2004a, b), indicate that the differences in the productivity distribution in the core region when compared with that in periphery cannot be simply characterized as the result of intensified competition. We might find possible clues to the differences in Marshallian externalities or urban externalities, which would mitigate market competition and allow small and low productivity firms to survive in core regions. While Syverson (2004a, b) emphasize the competition intensified by local market size, the difference from our results are rather natural because we cover all manufacturing products, which are largely traded across regions. By contrast Syverson (2004a) concentrates on an example of extremely localized competition (ready-mixed concrete) and Syverson (2004b) looks at the relations with transport costs and trade exposure. In the next section, we analyze how this cross-regional difference is explained by underlying economic geography factors, such as market potential.

Another finding to note is that  $\kappa$  becomes smaller over time in many regions. In particular,  $\kappa$  declines remarkably in many prefectures after the mid-1980s. Furthermore, the decline of  $\kappa$  in

---

<sup>12</sup> As shown in Appendix Table C, the basic patterns in the gamma distribution remain the same even if we include multi-plant firms.

periphery regions is substantial (e.g. from 0.7 to 0.5 in the Miyazaki and Nagasaki prefectures).<sup>13</sup> This might indicate that the impact of intensified market competition, which is likely to be accelerated by global competition and the development of domestic transportation networks in the 1980s, became more important in periphery than in core regions.

### 3.3. Relationships between distribution shapes and economic geography

Keeping our preliminary results on firm distributions in mind this section relates the estimates reported in the previous section with geographical variables in order to provide economic interpretations.

To investigate how economic geography affects firm productivity distributions, we estimate the following two equations. The dependent variables of the regressions are the shape parameter  $\kappa$  and the standard deviation  $\sigma$  for each prefecture; both are derived from the extended generalized gamma distribution in the last section.<sup>14</sup>

$$\kappa_{jt} = const + \alpha_1 \cdot MP_{jt} + \beta_1 \cdot Urban_{jt} + \gamma_1 \cdot KS_{jt} + \delta_1^Y \cdot YEAR_t + \delta_1^R \cdot PREF_j + \varepsilon_{1jt} \quad (4)$$

$$\sigma_{jt} = const + \alpha_2 \cdot MP_{jt} + \beta_2 \cdot Urban_{jt} + \gamma_2 \cdot KS_{jt} + \delta_2^Y \cdot YEAR_t + \delta_2^R \cdot PREF_j + \varepsilon_{2jt} \quad (5)$$

The prefecture is indexed by  $j$ , while the suffix  $t$  denotes the year. On the right-hand side of the regressions, the market potential,  $MP$ , is defined as in Harris (1954), that is:

$$MP_{jt} \equiv \sum_{n=1}^{47} \frac{GDP_{nt}}{D_{jnt}} \quad (6)$$

---

<sup>13</sup> The values for some prefectures fluctuate over time (e.g. between 0.3 and 0.5 in Chiba and between 0.3 and 0.6 in Kanagawa). Manufacturing clusters were formed in Kanagawa and Chiba in the 1970s and 80s due to good market access to central Tokyo. Some villages and towns in these prefectures experienced drastic transitions from agricultural to manufacturing areas. This might lead to time varying values of  $\kappa$ .

<sup>14</sup> We also use the average  $\mu$  derived from the same gamma distribution as another dependent variable. The estimation results of  $\mu$  are reported in Appendix Table D. As we have already discussed the average productivity gap between core and periphery, this section concentrates on higher moments, specifically  $\kappa$  and  $\sigma$ .

where  $D_{jn}$  is geographical distance of capitals between prefectures  $j$  and  $n$ .<sup>15</sup> As a measure of urbanization we include *Urban*, which is defined as the share of the population in Densely Inhabited Districts (DID) in each region.<sup>16</sup> To check the robustness of the estimates we also use the following alternative proxies of urbanization: GDP per capita, *Firm* (the total number of manufacturing firms), *Manufacturing* (the share of manufacturing in the region's GDP), and *Infra* (public capital stock for industrial use).<sup>17</sup> To control for cross-regional variations in industrial specialization, we include a value of the Krugman index, which is defined as

$$KS_{jt} \equiv \sum_i |s_{ijt} - s_{it}| \quad (7)$$

where  $s_{ijt}$  ( $s_{it}$ ) denotes the share of industry  $i$  in region  $j$  (in Japan) in total manufacturing employment.<sup>18</sup> This index takes the value of zero when the region's industrial structure is the same as the national average. While urbanization indices consider the region as a whole when capturing urban externality Krugman's specialized index focuses on how the region specializes in a particular industry or how a particular industry is concentrated in the region analyzed. Year dummies *YEAR* and prefecture dummies *PREF* are added in the fixed-effects model applied to our panel data, error terms are represented by  $\varepsilon$ .

Table 4 reports the FGLS panel estimation results. The shape parameter  $\kappa$ , in (4), is significantly negatively related to market potential but positively related to urbanization of the

---

<sup>15</sup> When  $j=n$ , the internal distance is calculated by  $\frac{2}{3} \sqrt{\frac{Area}{\pi}}$  where "Area" denotes area of the prefecture  $j$ . (See Combes and Overman, 2004)

<sup>16</sup> DID is defined by the district of which population density is more than 4,000 people per square kilometer and population in adjacent area is more than 5,000. The data is taken from the Population Census.

<sup>17</sup> The prefecture-specific data for GDP, population and infrastructure are taken from Fukao and Yue (2000)'s data set.

<sup>18</sup> The estimates  $\kappa$  and  $\sigma$  are region-specific but not industry-specific. However, we have confirmed in Figure 3 that our principal results on productivity distributions are not affected by cross-regional differences in industrial compositions. The addition of the Krugman index to our regression controls for the region's industrial specialization patterns (in deviation from the national average).

region and industrial localization. With respect to the standard deviation,  $\sigma$  in (5), the coefficients on market potential, industrial localization and urbanization indices are all significantly positive. These results are robust across alternative indices of urbanization.

Combined with the descriptive statistics reported in previous sections, these regression results are informative in interpreting the core-periphery contrast. Firstly the standard deviation of the productivity distribution tends to have a significantly wider dispersion in urbanized regions. The productivity dispersion is also wider in regions with stronger market potential and regions with localization of industries. These results imply that high average income and demand in urban areas appears to accommodate wide ranges of firms (in productivity but also possibly in differentiated varieties). On the other hand, poor periphery regions with small local demand can support only a narrow range of firms. This effect of agglomeration on  $\sigma$  has not been detected in previous studies, including Syverson (2004a, b).

Secondly the shape parameter  $\kappa$  in regions with stronger market potentials tends to be significantly lower. This indicates that competition intensified by strong market potential leads the productivity distribution to be relatively close to the normal distribution. This finding is in line with Cabral and Mata (2003) in that both discover that more intense market competition leads to productivity distribution which is closer to the normal distribution.

Thirdly we find that the shape parameter  $\kappa$  is positively related to the region's urbanization and industrial localization. This implies that urbanized regions, or regions with concentration of specific industries, can accommodate low-productivity firms along a long tail of a left-skewed distribution. Our focus on the shape parameter differentiates us from previous work neglecting higher moments of productivity distributions. Our previous descriptive finding in Table 8 that the value of  $\kappa$  in core regions is often low, but not extremely low in all such regions, is possibly due to two offsetting effects ( $\kappa$  is related positively to urbanization but



negatively to market potential). Our regression disentangles agglomeration economies (competition mitigated by general urbanization or by localization of specific industries) from competition intensification effects (competition intensified by market potential), both of which are likely to co-exist in core regions and have been unnoticed in previous research.

In sum, the core-periphery contrast is straightforward in the standard deviation  $\sigma$ , as core regions are often urbanized, and have stronger market potentials and higher concentrations of industries. Wider ranges of firms are active in core regions due to large local demand based on the region's high per-capita income, large local market size, or good access to surrounding markets. However, by considering the shape parameter  $\kappa$ , this paper unveils that competition tends to be particularly intense (productivity distribution close to the normal distribution) where the region has strong market potential but is not urbanized or has localization/concentration of no specific industry. This finding suggests that low-productivity firms should be forced to exit low-wage rural regions producing goods for export and is consistent with our observation that only highly productive firms can profitably operate in export-platform locations.

#### **4. Discussions**

We now discuss the implications of our results in relation to the established literature. While many previous studies have investigated the agglomeration premium, most of them use aggregate regional data. By contrast, we use micro-firm data. Thus firm heterogeneity is a central issue in our paper. Firm heterogeneity is taken into account in recent theoretical models of economic geography, for example Baldwin and Okubo (2006) show that in geographical selection and sorting of firms, high productivity firms are more profitable and footloose and thus locate in large and competitive markets. The intensive competition in large market leads to a selection of high productivity firms. Our finding of higher average productivity in core regions

is consistent with this prediction. However, there are some aspects not captured by these economic geography models. First, there appears no clear cut-off in spatial selection and exists a substantial overlap in firm productivity between core and periphery regions. Second, periphery regions, especially low-income regions with good access to neighboring markets, have tougher competition and productivity distributions are less left-skewed, this might result from trade cost reduction facilitated by the development of transport systems. Finally and much more importantly, we find a competition-mitigating effect such as a Marshallian externality in core regions, which leads to the survival of low productivity firms. As a related theoretical contribution, Okubo, Picard, and Thisse (2008) prove that, although productive (unproductive) firms choose to locate in large competitive (small less competitive) regions by spatial sorting, high-cost firms could also locate in large markets where markets have good access due to co-agglomeration. Our firm-level empirical findings are in line with such theoretical predictions and indicate that the relation between agglomeration and firm productivity distribution is more nuanced than that simply captured by the intensification of local competition.

## **5. Concluding remarks**

This paper empirically studies the distribution of firm productivity across regions and finds that the distribution is substantially heterogeneous across regions. Average productivity is sizably high in core regions and the distribution of firm productivity is left-skewed and is far from conforming to a log-normal distribution. The periphery region is likely to have low productivity firms due to less competition, although tougher competition in recent years has led to such firms being pushed out. The core region, by contrast, has two interacting forces, which have been neglected in previous work concentrating on means and standard deviations of productivity distributions. While the severe competition induced by stronger market potential makes

productivity distributions closer to log-normal distributions, the urban externality accommodates firms with wider ranges of productivity to survive within the same market. As the impact of agglomeration on firm distribution is a critical concern, for many producers and policy makers, comparable micro-data studies in other countries will be useful in the future.

### **Acknowledgement**

This research is partly financed by Grant-in-Aid for Scientific Research and Japan's Research Institute of Economy, Trade, and Industry (RIETI). The access to official micro-data is arranged by RIETI.

### **Reference**

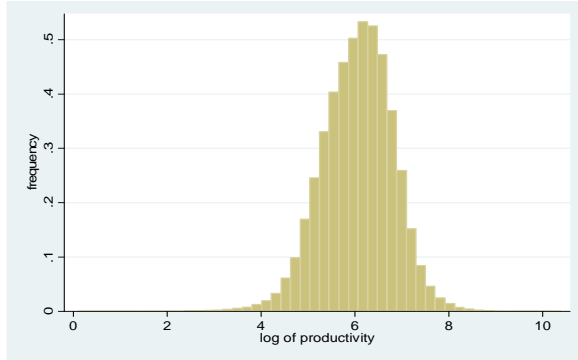
- Angelini, P and A. Generale (2005) "Firm size distribution: Do financial constraints explain it all? Evidence from survey data", *Economic working papers* 549, Bank of Italy, Economic Research Department.
- Angelini, P and A. Generale, (2008). "On the evolution of firm size distributions," *American Economic Review*, 98(1), pages 426-38
- Baldwin, R.E., and Okubo, T. (2006) "Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting," *Journal of Economic Geography* 6, 323-346.
- Barrios, S., Goerg, H., and Strobl, E. (2004). "The evolution of the firm size distribution and nationality of ownership," *Economics Bulletin*, 12(8), pages 1-7.
- Bernard, A., and Jones, C. (1996) "Comparing apples to oranges: productivity convergence and measurement across industries and countries," *American Economic Review* 86, 1216-1238.
- Brühlhart, M. and N. A. Mathys (2008) "Sectoral agglomeration economies in a panel of European regions," *Regional Science and Urban Economics* 38, 348-362.
- Cabral, L., and Mata, J. (2003) "On the evolution of the firm size distribution: facts and theory,"

- American Economic Review* 93, 1075-1090.
- Ciccone, A. (2002) "Agglomeration effects in Europe," *European Economic Review*, 46(2), 213-227.
- Combes, P.-Ph., Duranton, G., and L. Gobillon, L. (2008) "Spatial wage disparities: sorting matters!" *Journal of Urban Economics* 63, 723-742.
- Combes, P., and Overman, H. (2004) "The spatial distribution of economic activities in the European Union," In: Henderson, J., and Thisse, J. (Eds), *Handbook of Regional and Urban Economics*, Vol.4, Elsevier, pp. 2845-2909.
- Dekle, R. (2002) "Industrial concentration and regional growth: evidence from the prefectures," *Review of Economics and Statistics* 84, 310-315.
- Fujita, M., and Tabuchi, T. (1997) "Regional growth in postwar Japan," *Regional Science and Urban Economics* 27, 643-670.
- Fukao, K and Yue, X. (2000) "Sengo Nihon-kokunai ni okeru Keizaishusoku to Seisanyouso Tounyu—Solow seicho moderu ha tekiyou dekiru ka (Economic convergence and production input in post-war Japan—An application of Solow Economic Growth Model)", *Keizai Kenkyu* 52 (2).
- Harris, C. (1954) "The market as a factor in the localization of industry in the United States," *Annals of American Geographers* 64, 315-348.
- Henderson, V. (2003) "Marshall's scale economies," *Journal of Urban Economics* 53, 1-28.
- Holmes, T and J. Stevens, (2002). "Geographic concentration and establishment scale," *Review of Economics and Statistics*, 84(4), 682-690
- Melo, P., Graham, D., and Noland, R. (2009) "A meta-analysis of estimates of urban agglomeration economies," *Regional Science and Urban Economics* 39, 332-342.
- Nakamura, R. (2008) "Agglomeration effects on regional economic disparities: a comparison

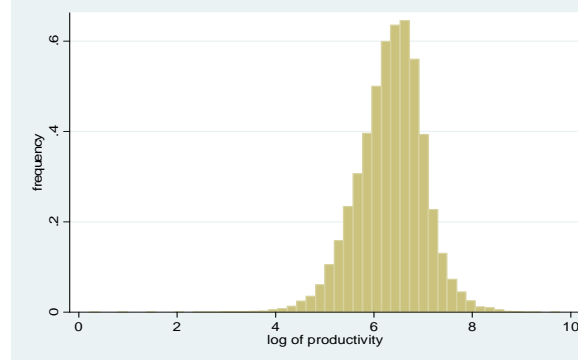
- between UK and Japan,” *Urban Studies* 45, 1947-1971.
- Okubo, T, Picard, P.M., and Thisse, J-F (2008) “The spatial selection of heterogeneous firms,” *CEPR Discussion Paper*, 6978.
- Rosenthal, S., and Strange, W. (2004) “Evidence on the nature and sources of agglomeration economies,” Chapter 49, in J. Henderson and J. Thisse (Eds) *Handbook of Urban and Regional Economics* vol.4, Elsevier, Amsterdam, pp.2119-2171.
- Strange, W. (2009) “Viewpoint: agglomeration research in the age of disaggregation,” *Canadian Journal of Economics* 42, 1-27.
- Syverson, C. (2004a) “Market structure and productivity: a concrete example,” *Journal of Political Economy* 112, 1181-1222.
- Syverson, C. (2004b) “Product substitutability and productivity dispersion,” *Review of Economics and Statistics* 86, 534-550.

Figure 1: Productivity Distribution (1990)

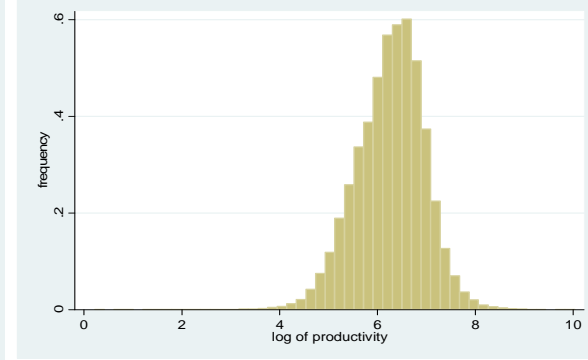
1-a: All regions



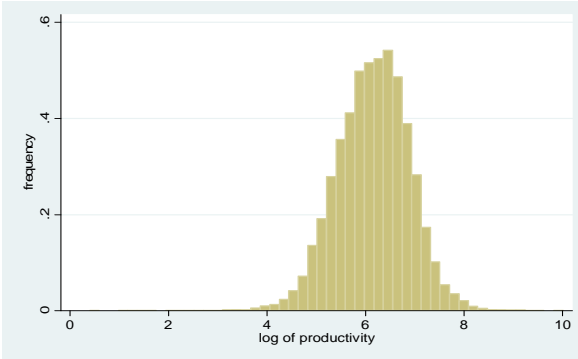
1-b: Tokyo



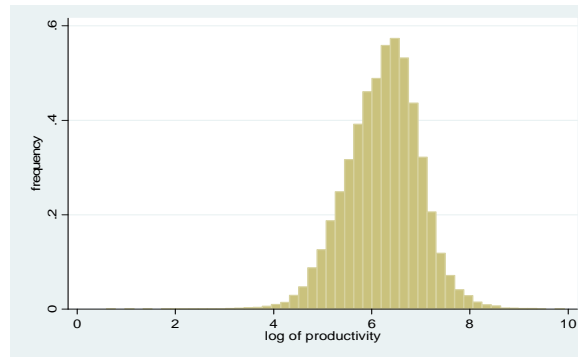
1-c: Greater Tokyo



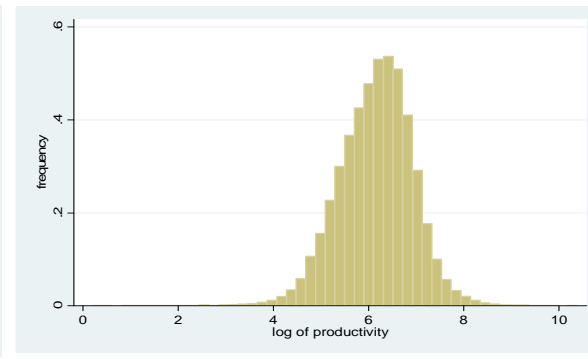
1-d: Aichi



1-e: Osaka



1-f: Greater Osaka



1-g: Core regions

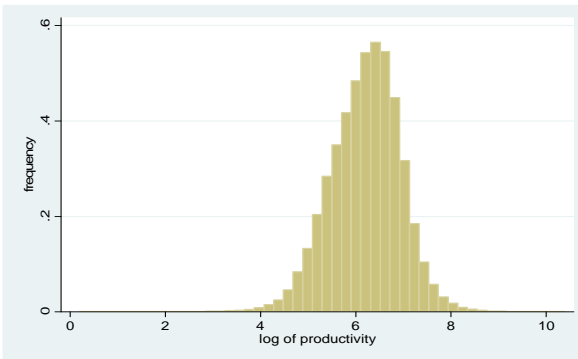
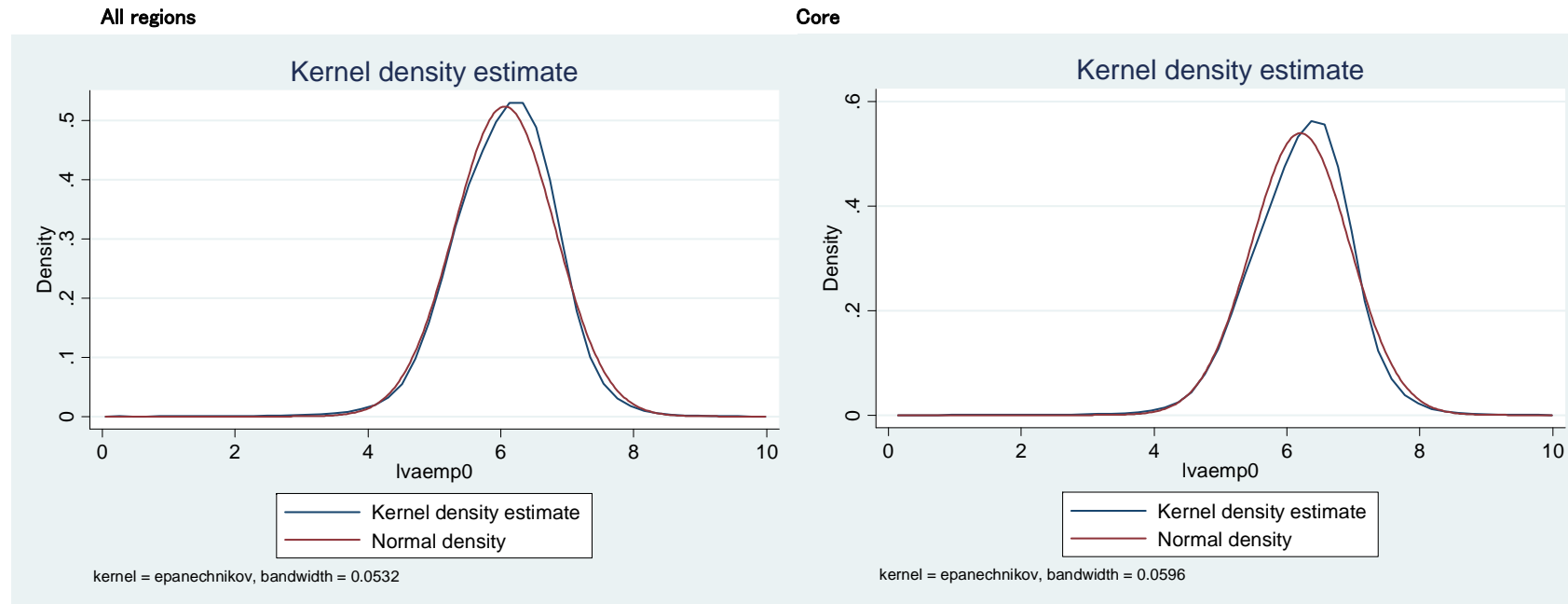


Figure 2: Kernel Density and Normal Distribution (1990)



Kernel density in Core and Periphery

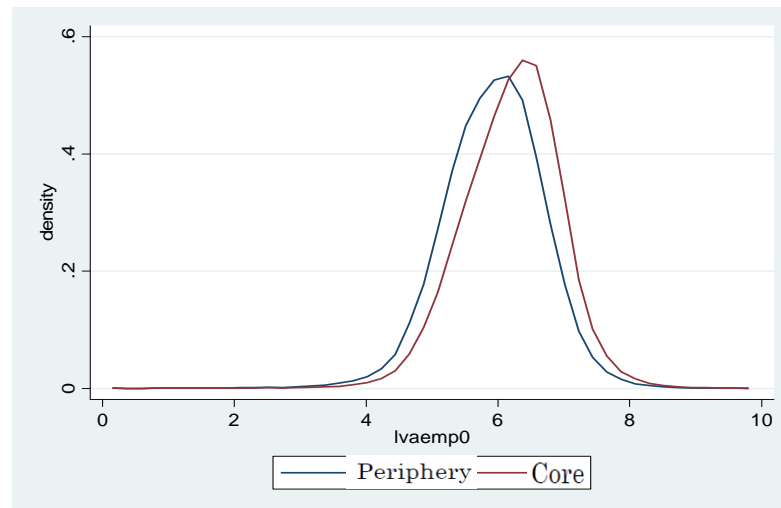
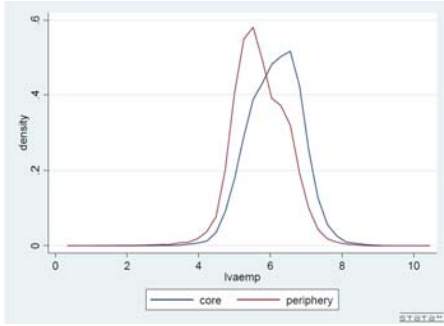
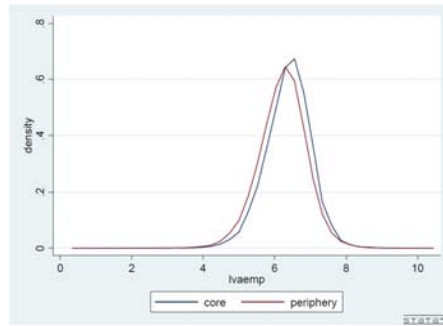


Figure 3: Kernel Density in Representative Sectors

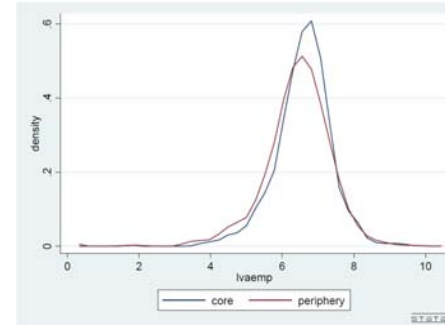
Electric machinery



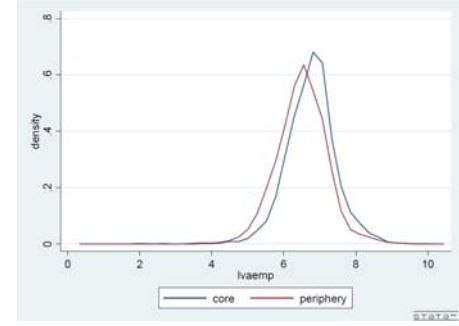
Metal



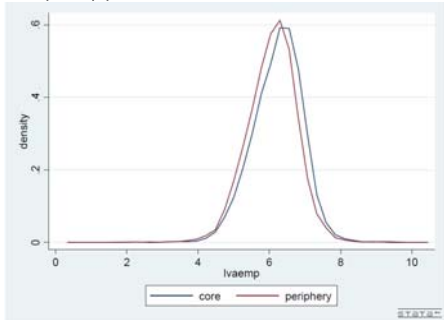
Chemical



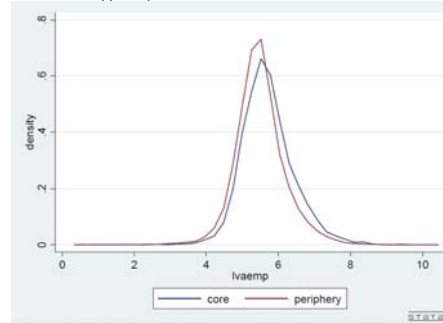
Iron and Steel



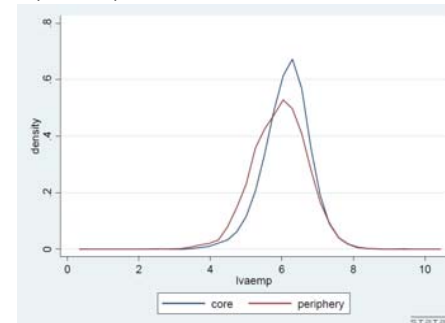
Transport equipment



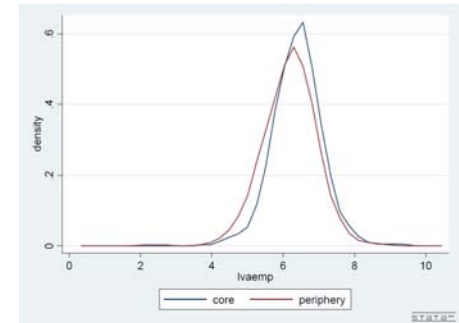
Textile and Apparel products



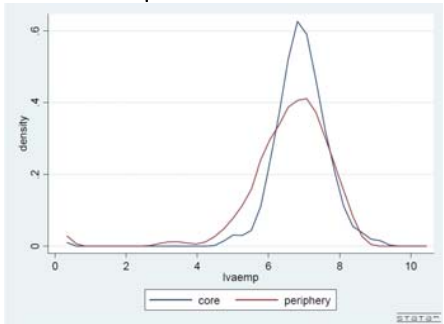
Paper and Pulp



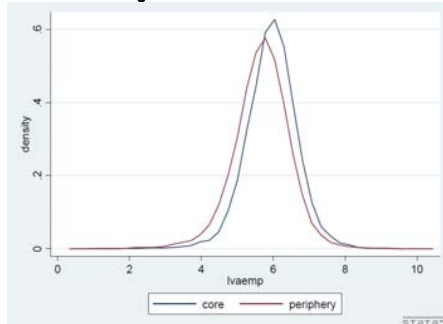
Nonferrous metal



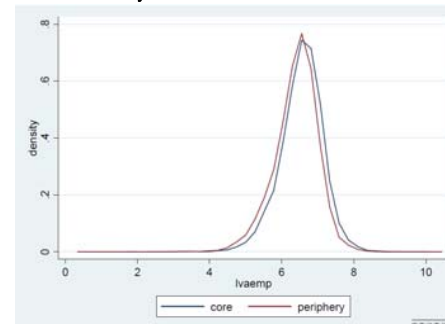
Petroleum and Coal products



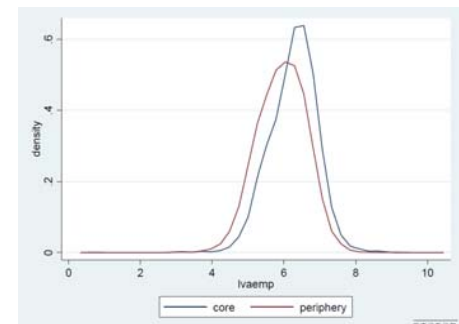
Food manufacturing



General machinery



Precision instrument





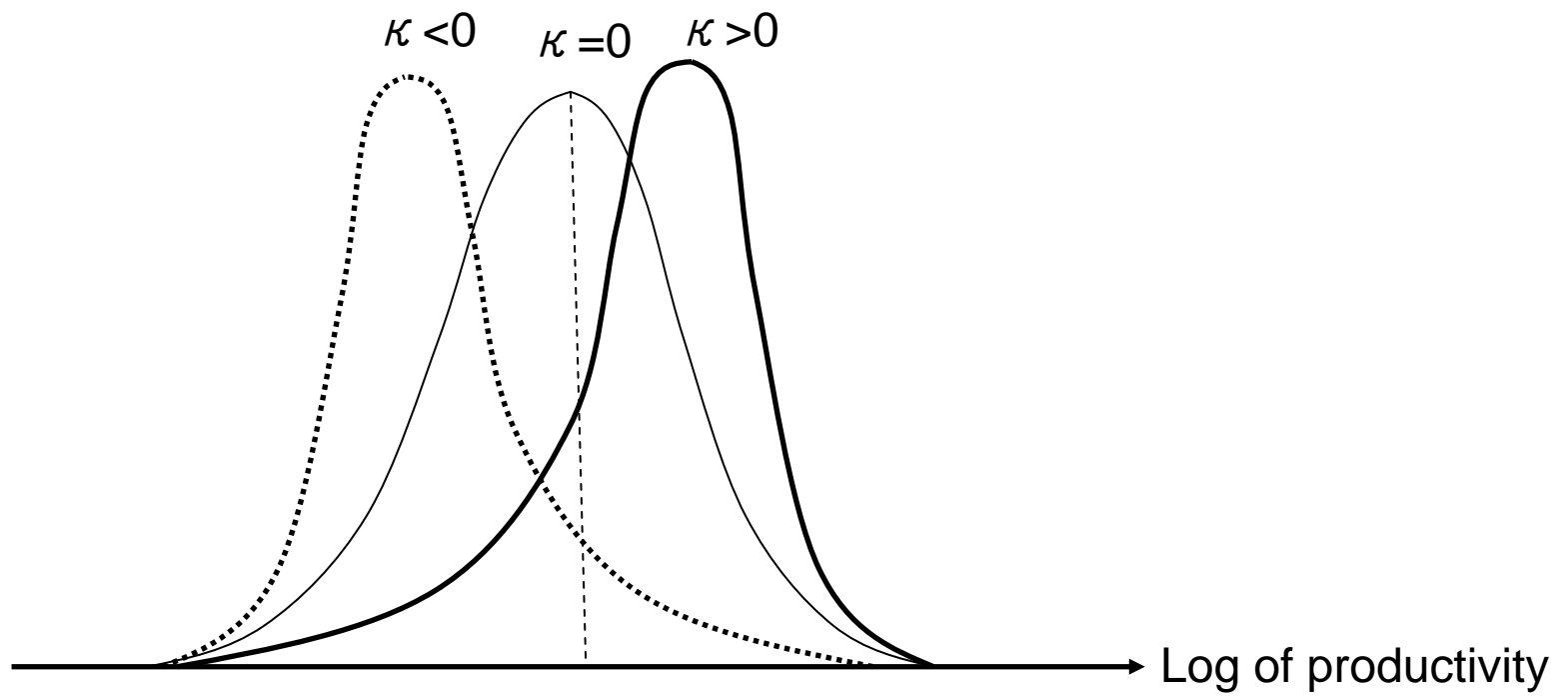
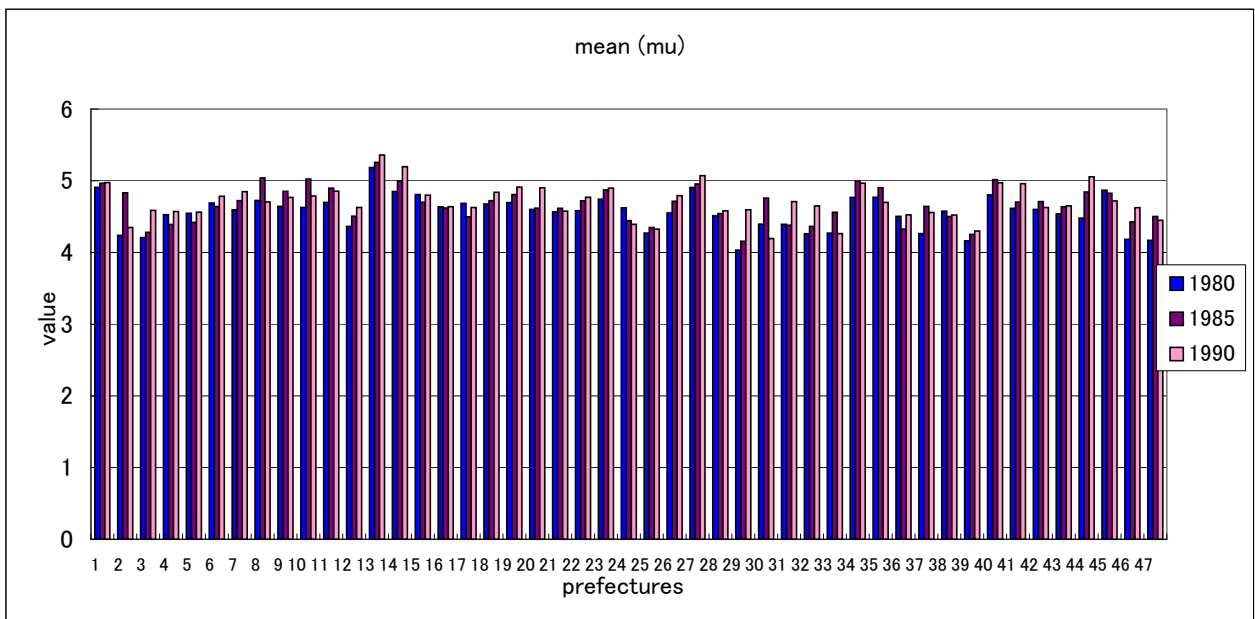
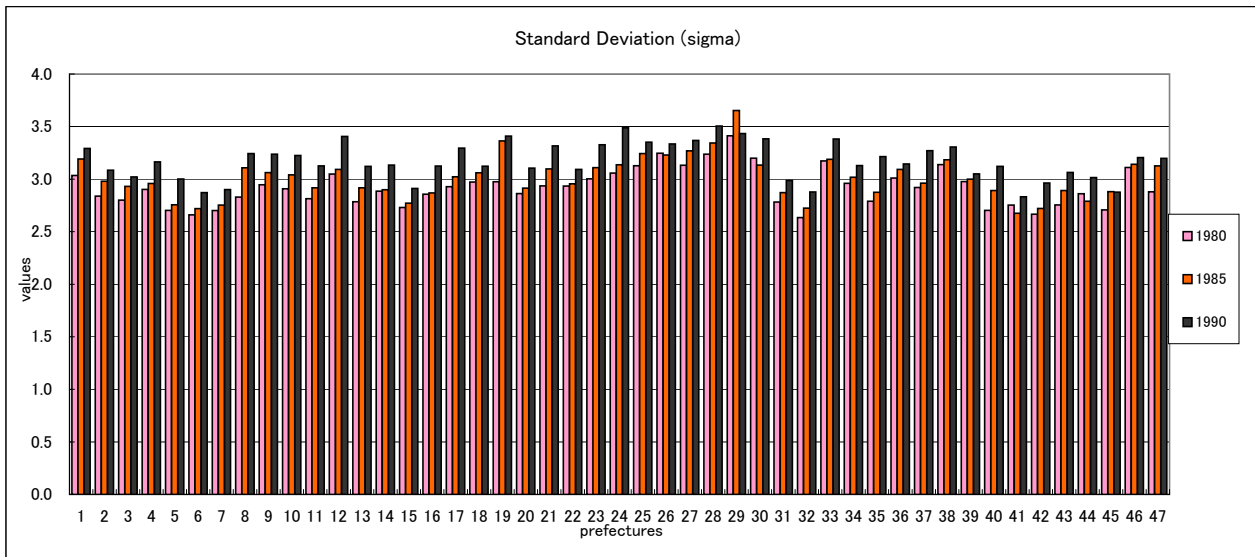
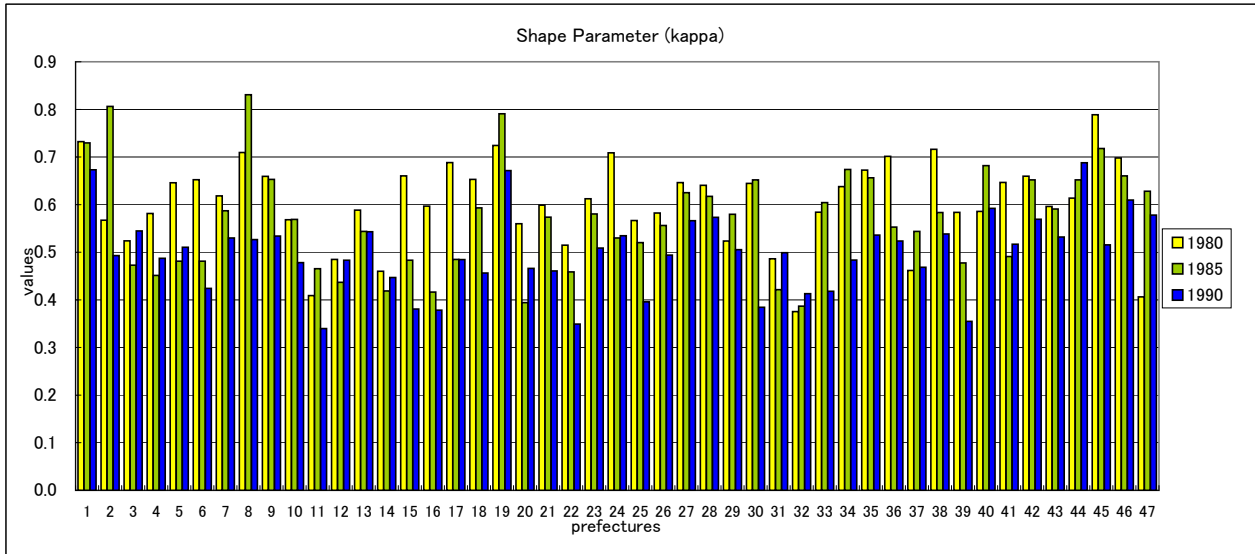


Figure 4: Gamma distribution and shape parameters ( $\kappa$ )

Figure 5: Gamma distribution estimates



**Table 1: Average Productivity Gap**

KS test

	1978	1980	1983	1985	1988	1990
Tokyo	0.1849	0.1907	0.1934	0.1945	0.1928	0.1932
Greater Tokyo	0.1585	0.1629	0.1589	0.1618	0.1567	0.1598
Greater Osaka	0.0932	0.092	0.0887	0.0778	0.0759	0.0683
Core	0.166	0.1685	0.1651	0.166	0.1607	0.1606

p-values for all results are all zero

**Table 2: Estimation of Productivity Premium**

year	1978			1980			1983			1985			1988			1990		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Tokyo	0.3783			0.442			0.4512			0.4886			0.4205			0.4402		
	[17.91]**			[20.97]**			[21.09]**			[22.01]**			[19.55]**			[19.67]**		
Osaka	0.339			0.3619			0.3531			0.3719			0.3081			0.3346		
	[14.47]**			[17.13]**			[16.48]**			[16.73]**			[14.33]**			[14.98]**		
Aichi	0.174			0.2273			0.2402			0.2838			0.23			0.2736		
	[8.17]**			[10.71]**			[11.15]**			[12.70]**			[10.65]**			[12.18]**		
Greater Tokyo	0.3184			0.4456			0.4482			0.2589			0.6326			0.4		
	[14.48]**			[17.96]**			[19.12]**			[10.40]**			[27.29]**			[15.61]**		
Greater Osaka	0.3398			0.2474			0.4416			0.2133			0.5893			0.3366		
	[15.23]**			[10.08]**			[18.60]**			[8.96]**			[24.88]**			[13.69]**		
Core		0.339			0.4315			0.4416			0.2589			0.3266			0.5337	
		[15.23]**			[17.17]**			[18.60]**			[10.40]**			[14.33]**			[21.04]**	
R-squared	0.106	0.106	0.106	0.106	0.115	0.115	0.111	0.111	0.111	0.119	0.119	0.119	0.137	0.137	0.137	0.137	0.137	0.137
F	385.63	385.63	385.63	385.63	387.89	387.89	385.4	385.4	385.4	361.44	361.44	361.44	315.03	315.03	315.03	315.03	315.03	315.03
Sample	348683			346333			355323			339814			332982			324687		

[ ]: t-values

\*\* : significant at 5%

\* : significant at 10%

All results are OLS regressions.

Prefecture dummies and sector dummies are included in all regressions.

Table 3: Gamma Distribution

Code	1978			1980			1983			1985			1988			1990		
	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$
Total	2.880	0.585	4.544	3.015	0.591	4.641	3.062	0.582	4.710	3.124	0.577	4.755	3.161	0.542	4.808	3.293	0.492	4.761
1	2.955	0.763	4.804	3.034	0.732	4.909	3.087	0.721	4.957	3.192	0.730	4.966	3.199	0.683	5.051	3.291	0.673	4.976
2	2.753	0.649	4.392	2.838	0.567	4.238	2.859	0.699	4.637	2.979	0.806	4.833	3.057	0.806	4.898	3.085	0.493	4.347
3	2.739	0.691	4.369	2.800	0.524	4.208	2.870	0.404	4.128	2.930	0.473	4.280	2.890	0.519	4.535	3.021	0.545	4.586
4	2.736	0.479	4.321	2.902	0.581	4.527	2.958	0.583	4.560	2.959	0.451	4.390	3.093	0.566	4.634	3.164	0.487	4.572
5	2.522	0.314	4.134	2.703	0.646	4.548	2.693	0.353	4.215	2.756	0.481	4.419	2.883	0.480	4.446	3.001	0.511	4.562
6	2.675	0.512	4.267	2.661	0.652	4.690	2.693	0.485	4.563	2.719	0.481	4.640	2.785	0.347	4.566	2.872	0.424	4.784
7	2.624	0.641	4.440	2.701	0.619	4.596	2.762	0.741	4.886	2.753	0.587	4.723	2.812	0.611	4.862	2.901	0.530	4.848
8	2.675	0.628	4.511	2.829	0.439	4.725	2.871	0.562	4.592	3.107	0.831	5.040	3.095	0.567	4.721	3.243	0.527	4.705
9	2.788	0.570	4.424	2.946	0.659	4.643	3.037	0.656	4.712	3.062	0.653	4.853	3.091	0.569	4.783	3.237	0.534	4.769
10	2.822	0.551	4.469	2.908	0.598	4.627	2.962	0.584	4.753	3.041	0.569	5.026	3.234	0.809	5.188	3.224	0.478	4.788
11	2.746	0.488	4.663	2.814	0.409	4.698	2.837	0.434	4.815	2.917	0.465	4.898	2.947	0.363	4.852	3.127	0.340	4.855
12	2.855	0.436	4.244	3.048	0.485	4.365	3.010	0.482	4.540	3.092	0.437	4.506	3.231	0.334	4.367	3.406	0.483	4.626
13	2.635	0.534	5.014	2.785	0.589	5.185	2.909	0.589	5.221	2.917	0.544	5.257	2.937	0.540	5.363	3.121	0.543	5.359
14	2.846	0.643	4.959	2.885	0.480	4.850	2.818	0.357	4.850	2.899	0.419	4.998	2.988	0.400	5.051	3.134	0.447	5.196
15	2.653	0.625	4.613	2.730	0.680	4.809	2.741	0.506	4.649	2.771	0.483	4.701	2.829	0.514	4.871	2.911	0.381	4.800
16	2.715	0.504	4.400	2.856	0.597	4.636	2.811	0.453	4.600	2.869	0.416	4.620	2.898	0.360	4.673	3.125	0.379	4.639
17	2.830	0.714	4.612	2.928	0.688	4.685	2.992	0.608	4.640	3.022	0.485	4.498	3.124	0.660	4.920	3.295	0.485	4.628
18	2.780	0.649	4.566	2.972	0.653	4.677	2.944	0.470	4.557	3.060	0.593	4.723	3.019	0.477	4.756	3.123	0.456	4.840
19	2.859	0.777	4.706	2.976	0.724	4.697	3.207	0.676	4.577	3.364	0.719	4.809	3.140	0.568	4.752	3.409	0.671	4.911
20	2.738	0.496	4.407	2.862	0.560	4.600	2.921	0.480	4.571	2.915	0.493	4.617	2.965	0.420	4.760	3.104	0.466	4.902
21	2.826	0.573	4.435	2.935	0.599	4.566	3.042	0.627	4.650	3.098	0.574	4.616	3.226	0.535	4.605	3.317	0.461	4.576
22	2.793	0.540	4.538	2.935	0.515	4.582	2.950	0.522	4.697	2.956	0.459	4.722	3.006	0.442	4.828	3.092	0.349	4.771
23	2.930	0.656	4.679	3.002	0.612	4.744	3.054	0.613	4.840	3.109	0.580	4.873	3.191	0.590	4.972	3.326	0.509	4.898
24	2.925	0.681	4.433	3.058	0.709	4.622	3.049	0.681	4.695	3.137	0.530	4.443	3.316	0.572	4.478	3.488	0.535	4.392
25	2.941	0.517	4.149	3.129	0.567	4.271	3.162	0.663	4.551	3.244	0.520	4.348	3.284	0.467	4.362	3.352	0.395	4.326
26	3.164	0.658	4.599	3.247	0.582	4.551	3.186	0.620	4.818	3.230	0.556	4.713	3.352	0.606	4.863	3.334	0.494	4.793
27	2.971	0.635	4.811	3.132	0.646	4.906	3.206	0.670	4.987	3.268	0.625	4.956	3.260	0.608	5.024	3.367	0.566	5.072
28	3.005	0.600	4.437	3.237	0.641	4.514	3.269	0.650	4.604	3.344	0.617	4.542	3.328	0.566	4.586	3.505	0.573	4.581
29	3.187	0.644	4.253	3.412	0.424	4.034	3.582	0.476	3.924	3.652	0.580	4.155	3.391	0.504	4.506	3.433	0.506	4.595
30	3.052	0.635	4.364	3.192	0.645	4.395	3.130	0.596	4.498	3.134	0.652	4.760	3.306	0.555	4.459	3.383	0.384	4.194
31	2.595	0.496	4.456	2.782	0.486	4.392	2.737	0.406	4.411	2.872	0.421	4.380	2.990	0.570	4.712	2.988	0.499	4.710
32	2.480	0.430	4.316	2.634	0.375	4.262	2.605	0.335	4.324	2.723	0.387	4.363	2.846	0.506	4.641	2.878	0.413	4.650
33	2.959	0.626	4.334	3.174	0.584	4.269	3.178	0.630	4.487	3.188	0.604	4.560	3.204	0.303	4.146	3.382	0.418	4.264
34	2.815	0.524	4.491	2.959	0.638	4.769	2.884	0.603	4.876	3.018	0.674	4.993	3.049	0.608	5.027	3.129	0.484	4.966
35	2.631	0.594	4.545	2.790	0.672	4.770	2.794	0.666	4.899	2.874	0.656	4.905	3.015	0.628	4.892	3.214	0.536	4.700
36	2.725	0.495	4.165	3.010	0.702	4.504	3.033	0.592	4.351	3.093	0.553	4.325	3.008	0.504	4.465	3.144	0.524	4.524
37	2.849	0.526	4.274	2.921	0.462	4.263	2.953	0.660	4.720	2.962	0.544	4.644	3.054	0.466	4.644	3.271	0.469	4.559
38	3.003	0.514	4.096	3.138	0.716	4.577	3.150	0.578	4.381	3.183	0.583	4.502	3.178	0.643	4.732	3.306	0.539	4.523
39	2.987	0.596	4.044	2.977	0.584	4.163	3.034	0.479	4.109	2.999	0.477	4.253	3.109	0.549	4.476	3.051	0.355	4.300
40	2.552	0.603	4.741	2.702	0.586	4.803	2.704	0.539	4.849	2.891	0.682	5.016	2.957	0.612	5.017	3.121	0.592	4.972
41	2.538	0.613	4.618	2.753	0.646	4.616	2.620	0.549	4.761	2.675	0.491	4.703	2.807	0.566	4.885	2.832	0.517	4.960
42	2.520	0.667	4.515	2.667	0.660	4.600	2.679	0.571	4.533	2.722	0.652	4.709	2.748	0.438	4.488	2.962	0.569	4.626
43	2.617	0.456	4.214	2.755	0.596	4.539	2.783	0.416	4.394	2.892	0.591	4.638	2.972	0.595	4.761	3.064	0.532	4.650
44	2.775	0.658	4.478	2.861	0.614	4.479	2.881	0.680	4.742	2.789	0.652	4.843	2.809	0.712	5.097	3.014	0.688	5.056
45	2.531	0.627	4.552	2.707	0.789	4.869	2.776	0.611	4.628	2.881	0.718	4.827	2.841	0.550	4.671	2.877	0.516	4.719
46	2.930	0.653	4.171	3.111	0.698	4.184	3.047	0.560	4.223	3.141	0.660	4.426	3.117	0.382	4.196	3.205	0.610	4.625
47	2.929	0.761	4.641	2.879	0.406	4.170	3.073	0.924	4.971	3.126	0.628	4.502	3.073	0.676	4.533	3.198	0.578	4.449
48	2.842	0.643	4.747	2.933	0.643	4.747	2.933	0.643	4.747	2.933	0.643	4.747	2.933	0.643	4.747	2.933	0.643	4.747

\*\* significant at 5%  
Bottom line in each estimate is -value.

Table 4: FGLS Results on Productivity Distribution Heterogeneity

	1	2	3	4	5	6	7	8
Dependent variables	$\kappa$	$\kappa$	$\kappa$	$\kappa$	$\sigma$	$\sigma$	$\sigma$	$\sigma$
MKT	-0.0761 [-7.62]**	-0.0564 [-5.92]**	-0.0616 [-6.14]**	-0.0546 [-5.07]**	0.11967 [8.74]**	0.10119 [5.66]**	0.11182 [6.29]**	0.1117 [8.08]**
KS	0.0961 [2.73]**	0.0903 [2.60]**	0.1371 [3.71]**	0.0676 [1.93]*	0.15951 [3.60]**	0.19513 [4.11]**	0.17544 [3.69]**	0.09059 [1.92]*
Urban	0.0839 [2.60]**			0.0983 [2.30]**	0.18495 [4.44]**		0.15916 [3.27]**	0.33173 [6.43]**
Firm	0.0333 [4.13]**					0.03901 [3.42]**	0.01354 [1.00]	
GDPcapita		0.0729 [2.61]**						
Manufacturing			0.0405 [6.41]**					
Infla		0.0322 [4.10]**		0.0231 [2.20]**				-0.06545 [-4.01]**
Wald Chi-2	175.63	172.62	140.74	142.42	725.47	714.14	732.49	772.36

constant term is omitted  
time dummies are omitted  
FGLS panel with heteroskedastic but uncorrelated error structure  
Number of observations is 282. Number of groups is 47.  
[ ]: z-values  
\*\* 5%      \* 10% significance

## Appendix

### Table A: Basic Statistic

#### Firm productivity (in logarithm)

	1978	1980	1983	1985	1988	1990
Obs	348683	346333	355323	339814	332982	324687
Mean	5.466605	5.617749	5.712944	5.79242	5.902257	6.025851
Std. Dev.	0.807285	0.831771	0.829244	0.848545	0.832791	0.87742
Variance	0.651709	0.691843	0.687645	0.720028	0.693541	0.769866
Skewness	-1.30789	-1.46736	-1.46885	-1.64182	-1.59998	-1.8487
Kurtosis	10.96468	11.97741	12.03113	12.98473	13.01716	14.1606
percentail						
1%	3.386809	3.461262	3.555348	3.555348	3.772761	3.713572
5%	4.265025	4.394449	4.493121	4.564348	4.688521	4.774913
10%	4.564348	4.701616	4.795791	4.864967	4.976734	5.081404
25%	5.01728	5.166164	5.253582	5.331107	5.4375	5.55511
50%	5.509388	5.664695	5.766131	5.849325	5.959071	6.0898
75%	5.966916	6.126869	6.222472	6.316391	6.425949	6.570683
90%	6.367157	6.531461	6.620586	6.709914	6.809388	6.96755
95%	6.622838	6.790097	6.878326	6.958528	7.054782	7.210966
99%	7.200612	7.388603	7.473702	7.538894	7.617444	7.76797

#### Regression Variables

Variables	Obs	Mean	Std.Dev	Min	Max
$\kappa$	282	0.561326	0.106034	0.303	0.924
$\sigma$	282	2.983004	0.214936	2.48	3.652
$\mu$	282	4.622996	0.257272	3.92427	5.362784
MKT	282	14.0643	0.632308	12.31976	15.91758
KS	282	0.529467	0.15348	0.242133	1.046072
urban	282	0.433502	0.180862	0.209756	0.969981
Firm	282	8.750307	0.824496	7.131699	10.90658
GDPcapita	282	0.801253	0.290592	0.121016	1.963593
Inf	282	14.01443	0.660407	12.65949	15.94146
GDP	282	15.26669	0.858726	13.71362	18.2519

**Table B: Japanese Prefecture Code**

Prefecture Code	Pref Name	Area Name	Core
1	Hokkaido		
2	Aomori		
3	Iwate		
4	Miyagi		
5	Akita		
6	Yamagata		
7	Fukushima		
8	Ibaraki		
9	Tochigi		
10	Gunma		
<b>11</b>	<b>Saitama</b>	Greater Tokyo	Core
<b>12</b>	<b>Chiba</b>	Greater Tokyo	Core
<b>13</b>	<b>Tokyo</b>	Greater Tokyo	Core
<b>14</b>	<b>Kanagawa</b>	Greater Tokyo	Core
15	Niigata		
16	Toyama		
17	Ishikawa		
18	Fukui		
19	Yamanashi		
20	Nagano		
21	Gifu		
22	Sizuoka		
<b>23</b>	<b>Aichi</b>		Core
24	Mie		
25	Shiga		
<b>26</b>	<b>Kyoto</b>	Greater Osaka	Core
<b>27</b>	<b>Osaka</b>	Greater Osaka	Core
<b>28</b>	<b>Hyougo</b>	Greater Osaka	Core
29	Nara		
30	Wakayama		
31	Tottori		
32	Shimane		
33	Okayama		
34	Hiroshima		
35	Yamaguchi		
36	Tokushima		
37	Kagawa		
38	Ehime		
39	Kouchi		
40	Fukuoka		
41	Saga		
42	Nagasaki		
43	Kumamoto		
44	Oita		
45	Miyazaki		
46	Kagoshima		
47	Okinawa		

Table C: Gamma Distribution for all firms (single plant and multi-plant)

	1978			1980			1983			1985			1988			1990		
	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$	$\sigma$	$K$	$\mu$
Total	3.118	0.687	4.615	3.276	0.693	4.695	3.339	0.697	4.780	3.407	0.691	4.825	3.493	0.717	4.964	3.632	0.681	4.936
	983.77***	258.70***	665.73***	1023.24***	258.20***	642.07***	1082.70***	267.54***	658.33***	1075.36***	269.93***	655.91***	1098.49***	282.39***	661.06***	1137.10***	270.16***	644.93***
1	3.413	0.799	4.848	3.582	0.818	4.787	3.647	0.793	4.802	3.845	0.691	4.472	3.896	0.630	4.486	3.923	0.620	4.446
	161.01***	42.91***	86.79***	163.04***	42.24***	81.37***	162.45***	40.40***	79.36***	160.55***	20.15***	62.92***	166.97***	20.15***	64.26***	169.37***	27.75***	65.74***
2	3.122	0.607	4.113	3.293	0.787	4.337	3.332	0.837	4.440	3.393	0.782	4.580	3.619	0.893	4.734	3.600	0.782	4.556
	74.08***	14.00***	39.41***	76.95***	20.99***	42.67***	77.68***	22.59***	44.34***	77.32***	20.33***	41.99***	80.55***	23.32***	39.51***	84.08***	21.39***	41.47***
3	3.020	0.708	4.256	3.120	0.651	4.243	3.197	0.468	4.004	3.255	0.588	4.282	3.370	0.647	4.424	3.413	0.708	4.620
	80.20***	21.89***	52.74***	83.42***	19.52***	50.78***	88.15***	12.76***	46.68***	90.40***	17.83***	50.74***	95.52***	21.15***	52.89***	100.45***	25.08***	57.51***
4	3.016	0.640	4.440	3.191	0.754	4.680	3.245	0.644	4.523	3.286	0.603	4.433	3.535	0.778	4.763	3.559	0.623	4.579
	99.85***	24.40***	69.22***	104.89***	25.89***	69.55***	108.79***	24.83***	66.48***	108.85***	23.11***	65.57***	118.64***	32.34***	66.23***	120.43***	24.28***	64.17***
5	2.687	0.491	4.235	2.951	0.716	4.493	2.977	0.672	4.469	3.125	0.769	4.637	3.250	0.808	4.780	3.287	0.694	4.660
	75.74***	14.74***	61.89***	79.87***	22.12***	58.95***	82.57***	21.72***	59.33***	85.68***	25.29***	57.61***	90.04***	28.27***	59.83***	94.83***	22.40***	55.63***
6	2.878	0.636	4.353	2.883	0.679	4.624	2.965	0.665	4.667	3.005	0.612	4.655	3.076	0.658	4.870	3.145	0.573	4.861
	85.39***	23.91***	70.14***	94.67***	25.85***	73.71***	99.90***	25.75***	74.04***	101.49***	23.84***	74.86***	105.65***	27.39***	79.06***	110.04***	23.05***	73.73***
7	2.935	0.715	4.413	3.069	0.799	4.720	3.115	0.842	4.895	3.159	0.786	4.863	3.279	0.800	4.974	3.353	0.741	4.977
	114.79***	30.76***	78.22***	119.50***	35.10***	80.11***	121.88***	37.87***	81.68***	125.53***	35.55***	81.68***	133.17***	37.83***	83.01***	139.82***	36.56***	86.12***
8	3.120	0.752	4.526	3.341	0.809	4.674	3.346	0.720	4.639	3.570	0.827	4.853	3.664	0.738	4.753	3.860	0.722	4.745
	132.91***	36.98***	84.76***	142.93***	40.88***	82.19***	151.49***	37.58***	86.44***	157.53***	43.85***	83.04***	165.76***	39.63***	82.27***	174.35***	39.48***	79.92***
9	3.066	0.732	4.577	3.302	0.779	4.691	3.414	0.795	4.784	3.427	0.751	4.853	3.517	0.706	4.818	3.665	0.713	4.857
	132.41***	37.63***	91.29***	138.43***	38.47***	82.59***	148.28***	41.72***	89.91***	149.57***	39.96***	87.32***	154.90***	37.24***	86.26***	161.35***	39.30***	87.32***
10	3.036	0.684	4.603	3.138	0.647	4.651	3.234	0.729	4.893	3.396	0.657	4.816	3.429	0.816	5.159	3.622	0.632	5.250
	141.05***	36.95***	88.89***	146.10***	34.09***	85.98***	153.15***	41.54***	102.68***	153.71***	51.89***	104.23***	165.62***	50.47***	108.89***	172.13***	51.89***	103.79***
11	3.021	0.614	4.794	3.088	0.574	4.884	3.169	0.652	5.062	3.211	0.636	5.120	3.305	0.665	5.243	3.468	0.620	5.322
	209.72***	48.63***	157.04***	215.89***	46.06***	157.22***	234.61***	58.64***	171.62***	240.31***	58.81***	176.11***	252.85***	63.69***	179.59***	265.17***	58.89***	171.93***
12	3.453	0.716	4.433	3.662	0.704	4.439	3.628	0.750	4.725	3.675	0.707	4.729	3.927	0.771	4.823	4.112	0.773	4.850
	154.07***	37.85***	81.23***	159.98***	36.35***	75.41***	165.90***	41.47***	84.81***	168.15***	38.38***	85.52***	180.79***	44.84***	83.34***	183.53***	43.84***	78.24***
13	2.772	0.656	5.168	2.913	0.667	5.279	3.042	0.662	5.308	3.048	0.622	5.354	3.073	0.652	5.520	3.247	0.648	5.518
	318.79***	91.11***	307.72***	329.40***	91.22***	294.49***	341.01***	88.42***	281.29***	342.83***	82.83***	280.62***	321.43***	84.95***	276.59***	326.85***	80.80***	253.73***
14	3.197	0.786	5.084	3.244	0.696	5.089	3.212	0.748	5.306	3.291	0.747	5.365	3.376	0.772	5.529	3.502	0.727	5.543
	192.93***	57.21***	135.43***	196.87***	50.92***	138.92***	206.08***	60.46***	154.47***	210.50***	60.82***	154.74***	217.70***	64.07***	158.27***	224.81***	60.28***	152.53***
15	2.801	0.704	4.675	2.941	0.793	4.928	2.957	0.658	4.774	2.978	0.589	4.767	3.074	0.634	4.942	3.153	0.570	4.956
	141.00***	41.87***	118.35***	148.41***	49.84***	121.71***	149.78***	38.24***	116.86***	151.07***	35.89***	117.45***	154.62***	37.14***	114.75***	161.85***	34.29***	116.89***
16	2.973	0.699	4.578	3.092	0.657	4.625	3.050	0.653	4.795	3.097	0.609	4.818	3.231	0.704	5.062	3.310	0.677	5.000
	94.83***	26.49***	69.95***	97.32***	22.77***	64.04***	101.50***	25.33***	74.04***	102.42***	23.75***	74.34***	108.93***	29.06***	78.49***	114.23***	28.09***	72.78***
17	2.962	0.689	4.507	3.086	0.694	4.626	3.130	0.623	4.608	3.187	0.572	4.570	3.313	0.690	4.878	3.516	0.659	4.842
	118.45***	30.34***	81.86***	123.63***	31.49***	82.54***	126.43***	28.16***	84.84***	128.90***	26.69***	83.59***	128.12***	33.11***	84.43***	137.02***	32.28***	81.44***
18	2.867	0.611	4.495	3.091	0.622	4.564	3.086	0.499	4.528	3.223	0.633	4.706	3.240	0.570	4.781	3.354	0.560	4.879
	99.15***	23.59***	74.88***	102.44***	21.81***	64.54***	108.26***	18.82***	71.15***	108.17***	24.69***	70.11***	107.95***	22.22***	71.93***	112.25***	21.78***	70.62***
19	3.021	0.755	4.605	3.173	0.679	4.527	3.503	0.723	4.479	3.625	0.780	4.656	3.459	0.666	4.769	3.696	0.672	4.767
	83.54***	24.41***	58.29***	89.50***	21.24***	54.79***	96.95***	23.26***	49.81***	102.52***	26.80***	51.58***	101.45***	22.58***	56.75***	108.24***	24.05***	55.65***
20	2.944	0.642	4.532	3.079	0.664	4.657	3.190	0.639	4.665	3.142	0.588	4.799	3.200	0.559	4.856	3.342	0.544	4.913
	142.92***	36.20***	100***	148.65***	37.65***	102.91***	157.44***	37.29***	104.84***	157.16***	34.97***	111.71***	159.12***	32.08***	107.39***	165.51***	32.16***	108.13***
21	3.014	0.635	4.451	3.154	0.658	4.563	3.252	0.690	4.658	3.308	0.645	4.657	3.472	0.556	4.516	3.569	0.537	4.581
	157.91***	38.10***	103.97***	169.18***	39.39***	105.57***	175.43***	43.53***	104.47***	179.90***	41.36***	109.72***	190.22***	33.99***	101.01***	195.94***	33.60***	102.39***
22	3.001	0.686	4.686	3.159	0.679	4.745	3.253	0.716	4.865	3.252	0.709	4.992	3.333	0.720	5.137	3.430	0.673	5.132
	192.70***	50.90***	139.83***	203.33***	50.89***	136.17***	214.44***	56.66***	141.23***	217.40***	59.05***	150.95***	223.84***	61.04***	150.33***	230.72***	57.57***	151.79***
23	3.124	0.700	4.700	3.239	0.696	4.799	3.306	0.689	4.869	3.373	0.687	4.963	3.458	0.699	5.066	3.586	0.638	5.032
	277.89***	73.43***	188.73***	286.34***	73.14***	186.49***	297.21***	73.30***	188***	302.85***	75.13***	192***	308.99***	76.64***	191.06***	322.18***	71.52***	190.21***
24	3.188	0.765	4.449	3.432	0.825	4.625	3.473	0.819	4.727	3.542	0.695	4.480	3.696	0.648	4.390	3.854	0.637	4.352
	126.45***	35.91***	71.59***	131.89***	37.51***	71.72***	138.10***	38.54***	74.78***	142.18***	32.59***	72.93***	148.68***	32.89***	69.24***	154.82***	30.89***	67.79***
25	3.442	0.746	4.311	3.700	0.725	4.245	3.777	0.741	4.409	3.990	0.774	4.395	4.100	0.819	4.565	4.246	0.806	4.595
	107.32***	28.50***	56.95***	114.24***	26.50***	49.89***	118.18***	28.85***	54.12***	121.24***	29.08***	49.09***	124.16***	31.96***	50.89***	129.23***	32.62***	51.32***
26	3.258	0.658	4.595	3.396	0.656	4.643	3.350	0.749	4.998	3.407	0.655	4.833	3.528	0.681	4.952	3.560	0.635	4.965
	168.05***	38.77***	101.54***	171.50***	37.55***	96.24***	172.04***	45.97***	108.11***	172.17***	39.							



**Table D: Average Productivity (FGLS panel)**

Dependent	1	2
	$\mu$	$\mu$
GDP	0.15155 [9.68]**	0.10289 [5.30]**
KS	0.14534 [2.66]**	0.21843 [4.02]**
Urban	0.16527 [2.88]**	0.16264 [2.80]**
GDPcapita		0.36416 [8.37]**
Wald Chi- <sup>2</sup>	882.64	2591.32

constant is omitted  
time dummies are omitted

Number of Observations is 282. Number of groups is 47.

t-values

\*\* 5%      \* 10% significance