# Center for Economic Institutions

# Working Paper Series

**No. 2010-11**

## *"Agglomeration or Selection? The Case of the Japanese Silk-Reeling Clusters, 1908–1915"*

Yutaka Arimoto, Kentaro Nakajima
and Tetsuji Okazaki

March 2011

Center for Economic Institutions
Working Paper Series

Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
http://cei.ier.hit-u.ac.jp/English/index.html
Tel:+81-42-580-8405/Fax:+81-42-580-8333

# Agglomeration or Selection?
# The Case of the Japanese Silk-Reeling Clusters, 1908–1915*

Yutaka Arimoto† Kentaro Nakajima‡ Tetsuji Okazaki§

February 22, 2011

**Abstract**

We examine two sources of productivity improvement in the specialized industrial clusters of the early twentieth century Japanese silk-reeling industry. Agglomeration improves the productivity of each plant through positive externalities, shifting plant-level productivity distribution to the right. Selection expels less productive plants through competition, truncating distribution on the left. We find no evidence confirming a right shift in the distribution in clusters or that agglomeration promotes faster productivity growth. Rather, the distribution in clusters was severely left truncated, even for younger plants. These findings imply that the plant-selection effect was the source of higher productivity in the Japanese silk-reeling clusters.

*Keywords:* Economic geography; Heterogenous firms; Industrial clusters; Productivity
*JEL classification:* R12; O18; L10

## 1 Introduction

Plants in industrial clusters are more productive than those located in non-clusters. Indeed, positive association between the spatial concentration of economic activities and productivity has been empirically confirmed in economic literature. For example, labor density is known to have a positive effect on productivity in the U.S. (Ciccone and Hall, 1996) and EU countries (Ciccone, 2002) at the regional level. This also holds true at the plant-level in the U.S. high-tech industry (Henderson, 2003). Excellent review of the existing studies on spatial concentration and productivity can be found in Rosenthal and Strange (2004) and Melo et al. (2009).

*We are grateful to Toshihiro Matsumura, Noriaki Matsushima, and Yasusada Murata for kindly providing invaluable comments and suggestions. We also thank Janet Hanter, Ryo Horii, Ryo Kambayashi, Daiji Kawaguchi, Takashi Kurosaki, Masaki Nakabayashi, Kaoru Sugihara, Tomohiro Machikita, Raffaele Paci, Ryuhei Wakasugi, Kazuhiro Yamamoto and participants of the meetings of Socio-Economic History Society at Hiroshima University, ARSC at Kushiro Public University of Economics, JEA at Kyoto University, NARSC at San Francisco, WEHC at Utrecht University, and seminars at Hitotsubashi University, KISER, Kyushu University, Nihon University, RIETI, Tohoku University, the University of Tokyo, and University of Tsukuba for comments and discussions. Authors gratefully acknowledge financial support from the Japan Society for the Promotion of Science (#21330064 and #22223003). Arimoto gratefully acknowledges financial support from the Japan Society for the Promotion of Science (#1810041). Nakajima gratefully acknowledges financial support from the Japan Society for the Promotion of Science (#21730181). The usual disclaimer applies.

†Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Tel: +81-42-580-8346. E-mail: arimotoy@ier.hit-u.ac.jp

‡Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan.Tel/Fax: +81-42-580-8369. E-mail: nakajima.kentaro@gmail.com

§Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: okazaki@e.u-tokyo.ac.jp

The higher productivity of plants in industrial clusters has long been explained through *agglomeration effects*, which refer to positive localized externalities: transferring knowledge and innovating ideas among densely agglomerated workers, alleviating matching through thick labor markets, or reducing transaction costs by transacting with proximate firms. These positive externalities directly improve plant-level productivities in the form of "bonuses" in the agglomeration. On the other hand, recent theoretical developments in spatial economics with heterogeneous firms have shed light on another channel of productivity improvement: plant-selection. That is, in clusters, the intensification of competition expels less productive plants, and consequently, only relatively productive plants survive. Thus, clusters have higher average regional productivity, even though they do not actually improve the productivity of each plant. This effect was initially proposed in the field of international trade and later introduced into economic geography (Melitz, 2003; Melitz and Ottaviano, 2008; Behrens et al., 2009; Baldwin and Okubo, 2006). Empirically, Syverson (2004) finds that higher plant density truncates productivity distribution in the lower tail in the ready-mixed concrete industry, implying that low-productivity producers were less likely to survive under heightened competition. In the field of international trade, Corcos et al. (2010) have identified the selection effect induced by trade policy.

Which of these two effects, agglomeration or selection, is more important in improving plant-level productivity in specialized industrial clusters? Recently, a pioneering work by Combes et al. (2009) distinguished the magnitude of the two effects and found that higher productivity in French *metropolitan areas* could be mostly explained by agglomeration effects. However, can productivity improvement in *specialized* industrial clusters also be explained by agglomeration effects? It should be noted that the agglomeration effects in turn can be classified into two subcategories depending on the sources of the externalities: industry specific externalities and externalities through various industries. Glaeser et al. (1992) termed the former category Marshall-Arrow-Romer (MAR) externalities, and the latter category is termed Jacobs externalities (Jacobs, 1969). Since Combes et al. (2009) focused on the urban metropolitan employment areas and applied labor density in all of the sectors as the source of externalities, the agglomeration effect detected by them was the mixed effect of both MAR and Jacobs externalities. On the other hand, specialized industrial clusters of plants within the *same* industry such as the high-tech cluster in the Silicon Valley (US) or automobile clusters in Detroit (US) and Toyota (Japan) are mainly representative of MAR type of agglomeration economies.

The purpose of this paper is identifying the source of productivity improvement effects in specialized industrial clusters. Particularly, we focus on the Japanese silk-reeling industry in the period from 1908 to 1915, which is characterized by a number of attractive features that are beneficial to our purpose. First, clusters were existent in this period: the Japanese silk-reeling industry formed huge clusters in the central and northeastern regions of Japan. These regions were mostly mountainous, peripheral areas with few plants other than the silk-reeling ones, which means that interactions between industries were limited. Second, besides the plants in clusters, there existed numerous silk-reeling plants across Japan, which allows us to use regional variations in the empirical analysis. Third, the silk-reeling industry produced a single homogenous product— raw silk— with similar equipment and technologies using a simple production process. Moreover, our dataset includes the information regarding physical output and not the value of raw-silk. These features allow us to estimate plant-level *physical* productivities more accurately than value-based productivities, which suffer from price and quality differences.

To distinguish between agglomeration and selection effects empirically, we follow the approach proposed by Combes et al. (2009), which examines the distribution of plant-level productivity. Combes et al. (2009) have developed a framework that nests selection and agglomeration by extending the model presented in Melitz and Ottaviano (2008) and introducing the agglomeration economies as in Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002). The model provides empirical predictions that enable a distinction between the two effects through an examination of the characteristics of productivity distributions. Intuitively, the agglomeration effect will shift the distribution to the right by improving the productivity of all plants in a region but keeping the shape of the distribution unchanged. On the other hand, the selection effect will left truncate the distribution at a higher productivity level by expelling less productive plants from the market. Combes et al. (2009) have found that productivity differences between French metropolitan areas are explained mostly by agglomeration. While the agglomeration effect found in Combes et al. (2009) can be considered as a mixed effect of MAR and Jacobs externalities, we study MAR externalities by focusing on specialized industrial clusters.

We also contribute to the literature by developing a model of agglomeration and selection through a competition among plants over input procurement, rather than output sales. The literature mostly relied on monopolistic competition in the output market to explain agglomeration or selection. However, silk-reeling plants in our study were price-takers where the price of raw silk was determined in the international market. Instead of competing in the output market, they aggressively competed over input procurement (cocoon) and labor forces (female workers). In order to take this feature into account, we modify Syverson's (2004) selection model over competition in the sales of homogeneous output to competition over input procurement. While Syverson (2004) relies on demand density as the source of selection, we show that the regional difference of entry cost can generate industrial clusters endogenously through selection. We then introduce an agglomeration economy into the model as in Combes et al. (2009), which allows us to nest both agglomeration and selection effects.

Besides the effort of distinguishing agglomeration and selection effects by focusing on productivity distribution, we investigate further by examining the timing when productivity improvements occur. If the agglomeration effect is in place, we should observe a higher productivity growth of plants in clusters *after* they start operation. On the other hand, the selection effect should have expelled the less productive plants and improved the average productivity *before* operation. By comparing the productivity growth rate and productivities of newly entered plants across clusters and non-clusters, we can distinguish the two effects from a different perspective.

Our main empirical finding is that selection improves productivity in clusters in the Japanese silk-reeling industry. We first confirmed that plants in clusters indeed had relatively higher productivity on average. We then examined the productivity distributions and found that the width of the distribution in clusters was narrower and more severely left truncated than non-clusters, which suggests selection. On the other hand, no clear evidence confirming the right-shift of the distribution in cluster was found. These evidences are consistent with plant-selection but not the agglomeration effect. We also examined whether the productivity growth rate was higher in clusters as implied by the agglomeration effect to find that the extent of agglomeration did not affect the productivity growth rate. Moreover, productivity distribution was severely left truncated in clusters even for younger plants. These results suggest that improvement of productivity took place before operation through selection rather than agglomeration economies after entry. Contrary to the results of Combes et al. (2009), which emphasize the role of the agglomeration effect in metropolitan areas, our results suggest the importance of the selection effect in specialized industrial clusters.

The rest of this paper is organized as follows. The next section overviews the Japanese silk-reeling industry in the period from 1908 to 1915. Section 3 provides a theoretical explanation of industrial clusters and plant-level productivities and proposes an idea for identifying the agglomeration and selection effects. Section 4 provides our main results on the identification of the plant-selection effect. Section 5 discusses further identification analysis by focusing on the timing of the cluster effects. Finally, Section 6 concludes the paper.

## 2 Overview of Japanese silk-reeling industry

The silk-reeling industry was one of the major industries in pre-war Japan. For example in 1908, it employed 24.4% of the total factory workers in Japan,[1] and its product, raw silk, occupied 26.6% of the total export (Ministry of International Trade and Industry, 1962; Toyo Keiizai Shinposha, 1927). A distinctive feature of the silk-reeling industry was that it was composed of numerous small and medium-sized plants. For example, in 1908, Japan had more than 3,200 silk-reeling plants. Even the largest plant occupied merely 4.3% of the total silk production in 1908, and the median value of the market share was 0.058% (Ministry of Agriculture and Commerce, 1910). In this sense, the market structure of the silk-reeling industry was very competitive.

These silk-reeling plants formed several clusters, of which Nagano, Aichi and Gifu in the central region of Japan, were the largest. Indeed, 37.5% of the silk-reeling plants in Japan were located in these three prefectures in 1908. Of these clusters, the cluster in Suwa County in Nagano Prefecture was the largest (Ministry of Agriculture and Commerce 1910). In 1930, in order to edit the History of Hirano Village, the center of the Suwa silk-reeling cluster, the assigned editors conducted a survey for the major plant owners there on the reasons behind the development of the silk-reeling industry in the village: one of the most common answers was that they could not secure their living only through agriculture as their land holdings were small and the soil was not fertile. Some respondents also cited the lack of good alternative occupations as a reason (Hirano Village Office 1932, pp.560–562). These answers imply that the opportunity cost for entering the silk-reeling industry was lower owing to the natural conditions in Suwa District. These first nature causes facilitated silk-reeling start-ups. It is also notable that besides the large clusters, silk-reeling plants operated in other areas as well, that is, in non-clusters. Figure 1 is the map of Japan, indicating the density of silk-reeling plants in 1908. The dark colored areas represent the prefectures where silk-reeling plants were densely located.

= Figure 1 =

Of the many studies on the development of the silk-reeling industry in Japan, Ishii (1972) and Nakabayashi (2003) are basic references. The silk-reeling industry emerged in Japan in the Tokugawa Era, but its growth was accelerated by the opening of the economy in 1859. Under the free trade regime, the export of raw silk experienced a boom. Initially, raw silk was produced by the traditional hand-reeling technology (*zaguri*-reeling), but in the 1870s, a new technology, machine-reeling (*kikai*-reeling), was developed, which modified the imported European technology. While hand-reeling production stagnated owing to competition with Chinese products in the 1870s, machine-reeling production experienced growth in this period (Nakabayashi 2003, pp.66–68). Machine-reeling production exceeded hand-reeling production in 1894, and the latter witnessed a decline in 1900.

According to Nakabayashi (2003), the basic market condition was factored by the growth of the silk weaving industry in the U.S. The U.S. silk-weaving industry introduced its mass production system in the 1860s and preferred homogeneous raw silk in large lots; this was a challenge for the traditional silk-reeling industries in both Japan and China. Based on the machine-reeling technology, the emerging silk-reeling industry in Japan met the demand of the U.S silk-weaving industry and thereby grew rapidly.[2] In the U.S. market, each silk-reeling plant

[1] The denominator is the number of total factory workers in 1909 (Ministry of International Trade and Industry, 1962).

[2] For example, in 1908, the ratio of the export to production was 98.9% with respect to machine-reeled raw silk and the share of the U.S. in the total Japanese raw silk export was 74.0% (calculated from the statistics in Nakabayashi, 2003, pp.468–462).

in Japan, which as mentioned above, was very small, was basically a price-taker, and it could sell as much raw silk as it wanted to at the market price in Yokohama, the main exporting port.

According to Duran (1913), the typical machine-reeling process in that period was as follows: A silk-reeling plant purchased cocoons from sericulture peasants. Cocoons were boiled to unwind the cocoon filaments. Then, from a group of boiled cocoons, unwound filaments were bundled and reeled onto a small moving reel which was powered by water, steam, electricity, or gas. Young female workers played a key role in this production process. Each female worker was in charge of one reeling machine, and her ability and level of effort substantially affected the productivity and quality of the raw silk.

Owing to its relatively simple production process, the management, productivity, and survival of a silk-reeling plant essentially depended on two factors: cocoons and workers. However, congestion and fierce competition characterized the procurement of these factors.

In the cocoons market, competition for procurement was very severe in clusters, which raised cocoon prices not only in clusters but also in the adjacent prefectures or prefectures farther away (Hirano, 1990; Ishii, 1972, ch.4; Matsumura, 1992). The competition was amplified by the temporal and technological constraint that raw cocoons had to be transported quickly or dried appropriately to ensure that their quality and condition was maintained, since they perished in moist environments. Drying was also necessary to prevent the metamorphosis of the pupa. To relax competition, silk-reeling plants attempted to arrange cartels for joint purchase of cocoons and production reduction, or they engaged in vertical integration (tokuyaku torihiki) in the form of contract farming by concluding direct prior agreements with cocoon farmers with regard to price and quantity.

Most of the workers were young women. They were primarily recruited from surrounding areas so that they could easily commute from their homes. However, the enormous labor demand in clusters forced plants to hire from remote areas and, therefore, pay a fixed cost for boarding, food, and even education, which consequently increased the unit labor cost. Moreover, since silk reeling required some experience and skill, recruiting and training new workers was expensive for the plants. Therefore, poaching of trained workers from other factories was prevalent in clusters, and it gradually proliferated throughout the country (Kambayashi, 2001; Nakabayashi, 2003; Tojo, 1990).

In the silk-reeling industry, the performance-based wage system was widely used to provide workers with incentives to improve productivity as well as product quality by the early twentieth century. A distinctive characteristic of this system was that the wage of each worker was determined on the basis of her relative performance to the average of all the workers in the same plant. The performance was evaluated by several measures including labor productivity, cocoon productivity and product quality, which were strategic for the competitiveness and profitability of the plant. Nakabayashi (2003) reports the basic statistics of the worker-level wage at a plant in Suwa from the late nineteenth century to the early twentieth century. For example, sample mean, median, variance, and maximum and minimum values of wage per day in 1908 were 0.24, 0.25, 1.09, 0.62, and 0.018 yen, respectively (p.259). On the other hand, the average daily wage of female agricultural workers was 0.23 yen in the same year (Statistics Bureau, Management and Coordination Agency 1988, p.228). These facts imply that basically an unskilled female silk-reeling worker earned approximately the same amount as an average agricultural worker, but highly skilled silk-reeling workers in clusters could earn higher wages depending on their productivity.

It is notable that this wage system was first devised by Nakayama Co. in Suwa to diffuse within the district, and then was transmitted to other districts (Ishii, 1972, pp.291–307; Nakabayshi, 2003, pp.241–277). This diffusion process can be regarded as a case of rapid knowledge spillover in an industrial cluster.

## 3 Theoretical model and empirical strategy

We first develop a theoretical framework to model plant-selection in the context of the Japanese silk-reeling industry by modifying Syverson's (2004) selection model. Then, we incorporate the agglomeration effect.

### 3.1 Market structure

We consider the entry and production decisions of silk-reeling plants that procure cocoons from farmers, reel silk, and sell the final product. Output is sold in the export market (Yokohama) at an exogenous price $p$ set by the international market. Plants are price-takers and they can sell as much as they wish to in Yokohama.

We assume that the production of $q_i$ silk by plant $i$ entails labor cost $h_i q_i$, input purchase cost $w(Q)q_i$ for cocoons, and fixed cost $f$. Silk production relied heavily on female workers who reeled unwound cocoon filaments, and their skill and effort were crucial determinants of plant-level productivity. We denote $h_i$ as the effective labor required to produce one unit of output (raw silk). This variable can be considered as the plant's productivity: higher $h_i$ implies greater labor requirement and, hence, lower productivity. We normalize the cost of unit effective labor to unity.

The input purchase price $w(Q)$ can be considered as an aggregate inverse supply function, where $Q = \sum q_i$ is the total output in the region (or, in other words, factor demand). We assume that $w(Q)$ increases with $Q$ because greater silk production requires a higher demand for cocoons. Therefore, unlike Syverson (2004), where demand density is the key variable of focus, plants in our model do not compete over output sales but rather over input purchase (cocoons).

The profit of a plant can be represented as

$$\pi_i = pq_i - h_i q_i - w(Q)q_i - f$$

For simplicity, we assume the linear marginal input purchase price as $w(Q) = wQ$:

$$\pi_i = pq_i - h_i q_i - wQq_i - f. \tag{1}$$

This model has two stages. In stage one, each potential plant decides whether to pay a sunk entry cost $s$ to enter the market. After payment, a plant draws the labor unit requirement $h_i$ from distribution $g(h)$ with support $[0, \overline{h}]$, where $\overline{h}$ is an arbitrary upper bound. In stage two, each plant decides the level of production $q_i$ given one's own productivity parameter $h_i$ and forming an expectation of the total output in a region, $E(Q)$.

The expected profit of an entrant plant at stage two is

$$E(\pi_i) = (p - h_i)q_i - wE(Q)q_i - f. \tag{2}$$

The first-order condition with respect to $q_i$ is

$$\frac{\partial E(\pi_i)}{\partial q_i} = p - h_i - wq_i - wE(Q) = 0 \tag{3}$$

so the optimal output of a plant with marginal cost $h_i$ is

$$q_i{}^*(h_i) = \frac{p - h_i - wE(Q)}{w}. \tag{4}$$

Inserting $q_i{}^*$ back into the expected profit yields

$$E[\pi_i(h_i)] = \frac{[p - h_i - wE(Q)]^2}{w} - f. \tag{5}$$

The expected profit decreases with $h_i$. Therefore, a critical labor unit requirement draw $\hat{h}$ exists such that entrants drawing $h_i > \hat{h}$ choose not to produce. This cutoff labor unit requirement draw can be solved by setting $E(\pi_i) = 0$:

$$\hat{h}(E(Q)) = p - wE(Q) - \sqrt{wf}. \tag{6}$$

Inserting $p - wE(Q) = \hat{h} + \sqrt{wf}$ obtained from (6) into $E(\pi_i)$ yields operating profits, conditional on $h_i \leq \hat{h}$:

$$E\left[\pi_i(h_i|h_i \leq \hat{h})\right] = \frac{\left(\hat{h} - h_i + \sqrt{wf}\right)^2}{w} - f. \tag{7}$$

We assume a free-entry condition so that plants enter the market (i.e., pay $s$ and draw $c_i$) until the expected value of entry is equal to zero:

$$V^e = \int_0^{\hat{h}} \left[\frac{\left(\hat{h} - h + \sqrt{wf}\right)^2}{w} - f\right] g(h)dh - s = 0 \tag{8}$$

The equilibrium $\hat{h}$ is the value that solves this expression and it is a function of $g(h)$ and parameters $w$, $f$, and $s$.

## 3.2 Entry cost, selection effect, and emergence of clusters

We now consider regions that are symmetric, except for entry cost $s$. On the basis of the historical information in Section 2, we assume that $s$ varied across regions owing to the first nature or opportunity costs of starting-up a new business. We show that a region with lower $s$ imposes fiercer competition (i.e., lower cutoff unit labor requirement) but retains more producing plants, resulting in the formation of a cluster.

We first consider the effect of entry cost $s$ on the cut-off marginal unit labor requirement $\hat{h}$. The implicit function theorem implies that

$$\frac{d\hat{h}}{ds} = \frac{-(\partial V^e/\partial s)}{\partial V^e/\partial \hat{h}} = \frac{w}{2\int_0^{\hat{h}} \left(\hat{h} - h + \sqrt{wf}\right) g(h)dh} > 0. \tag{9}$$

**Implication 1** (*Selection effect*)**.** Regions with a lower start-up cost have lower cutoff labor unit requirement $\hat{h}$. This implies that competition is more severe: high-cost (low-productivity) plants are not profitable in regions with a lower start-up cost.

Next, we investigate the effect of $s$ on the number of plants in a region. Let $N_e$ denote the number of *entrant plants* that had paid $s$ and drawn $h_i$, and let $N_p$ denote the number of *producing plants* with $h_i \geq \hat{h}$. Then,

$$N_p = N_e \int_0^{\hat{h}} g(h)dh = N_e G(\hat{h}). \tag{10}$$

By applying (4), the total production in a region can be represented as

$$Q = N_e \int_0^{\hat{h}} q_i^* g(h)dh = N_e \int_0^{\hat{h}} \frac{p-h}{w} g(h)dh - N_e E(Q)G(\hat{h}).$$

Since $Q = E(Q)$ in the equilibrium, we can represent $Q$ as a function of $N_e$ and $\hat{h}$:

$$Q(N_e, \hat{h}) = \frac{N_e}{1 + N_e G(\hat{h})} \int_0^{\hat{h}} \frac{p-h}{w} g(h) dh. \tag{11}$$

Additionally, rearranging (6) yields

$$Q = \frac{p - \hat{h} - \sqrt{wf}}{w}. \tag{12}$$

By applying (11) and (12), we can represent $N_e$ as a function of $\hat{h}$ and $g(\cdot)$ and parameters $\{p, w, f\}$:

$$N_e = \frac{p - \hat{h} - \sqrt{wf}}{\int_0^{\hat{h}} (\hat{h} - h + \sqrt{wf}) g(h) dh}. \tag{13}$$

Partial differentiation with respect to $\hat{h}$ yields

$$\frac{\partial N_e}{\partial \hat{h}} = \frac{-\int_0^{\hat{h}} (p-h) g(h) dh - (p - \hat{h} - \sqrt{wf})\sqrt{wf} g(\hat{h})}{\left[\int_0^{\hat{h}} (\hat{h} - h + \sqrt{wf}) g(h) dh\right]^2} < 0. \tag{14}$$

Therefore, lower $\hat{h}$ (fiercer selection) entails more entry.

Now, since $N_p = N_e G(\hat{h})$, the effect of the entry cost on the number of producing plants is expressed by

$$\frac{dN_p}{ds} = \frac{dN_e}{d\hat{h}} \frac{d\hat{h}}{ds} G(\hat{h}) + N_e g(\hat{h}) \frac{d\hat{h}}{ds}. \tag{15}$$

The effect of $s$ on $N_p$ consists of two effects. The first effect represented by the first term is the *entry effect*, which is negative. That is, *lower $s$* entails that more plants will enter the market. The second effect represented by the second term is the *competition effect*, which is positive since $\frac{d\hat{h}}{ds} > 0$ from (9). Because lower $s$ induces severe competition, the number of plants that can produce after entry will be smaller. The aggregate effect of a low entry cost on the number of producing plants depends on the relative magnitudes of the (positive) entry effect and the (negative) competition effect. However, we can show that the entry effect always dominates the competition effect. Substituting (13) and (14) into (15) yields

$$\begin{aligned} \frac{dN_p}{ds} &= \left[\frac{dN_e}{d\hat{h}} G(\hat{h}) + N_e g(\hat{h})\right] \frac{d\hat{h}}{ds} \\ &= \frac{-(p - \hat{h})\left[G(\hat{h})^2 - g(\hat{h}) \int_0^{\hat{h}} G(h) dh\right] - \left[G(\hat{h}) + \sqrt{wf} g(\hat{h})\right] \int_0^{\hat{h}} G(h) dh}{\left[\int_0^{\hat{h}} (\hat{h} - h + \sqrt{wf}) g(h) dh\right]^2} \frac{d\hat{h}}{ds} < 0. \end{aligned} \tag{16}$$

This is negative because $\left[G(\hat{h})^2 - g(\hat{h}) \int_0^{\hat{h}} G(h) dh\right]$ is positive since $G(\hat{h}) \geq G(h)$ for all $h \in [0, \hat{h}]$, $G(\hat{h}) > g(\hat{h})$, and $\frac{d\hat{h}}{ds} > 0$ from (9).

**Implication 2** (*Endogenous clusters*)**.** Lower entry cost $s$ entails that more plants will enter the market (entry effect), but it imposes fiercer competition after entry and reduces the number of producing plants (competition effect). The former always dominates the latter, and therefore, the number of producing plants is greater in a region with lower $s$.

### 3.3 Agglomeration effects and productivity distributions

Following Combes et al. (2009), we now introduce the agglomeration effect, which improves plants' productivities through the interaction between adjacent operating plants. We model this effect by assuming that when a plant interacts with $N_p$ plants, the effective units of labor supplied by an individual worker during their unit time becomes $a(N_p)$, where $a(0) = 1$, $a' > 0$, and $a'' < 0$. The improved productivity of workers should be compensated for by higher wages if the labor market is perfect.[3] However, a plant with unit labor requirement $h_i$ reduces the number of workers to $l(h_i) = q_i h_i / a(N_p)$ at a total cost of $a(N_p)l(h_i) = q_i h_i$. Thus, each plant's maximization problem is unchanged.

Given this agglomeration effect, the logarithm of a plant's productivity $\phi_i$ can be derived as follows:

$$\phi_i = \ln\left[\frac{q_i}{l}\right] = \ln\left[\frac{q_i}{q_i h_i / a(N_p)}\right] = \ln[a(N_p)] - \ln(h_i) \tag{17}$$

Then, the density function of the log productivities is as follows:

$$f(\phi) = \begin{cases} 0 & \text{for} \quad \phi < \hat{\phi} \equiv A - \ln(\hat{h}), \\ \frac{e^{A-\phi}g(e^{A-\phi})}{G(\hat{h})} & \text{for} \quad \phi \geqslant \hat{\phi}, \end{cases} \tag{18}$$

where $A \equiv \ln[a(N_p)]$.

As discussed above, each plants' the maximization problem is unchanged regardless of the presence of the agglomeration effect. Thus, plugging the equilibrium cut-off unit labor requirement $\hat{h}$ obtained from eq. (8) to eq. (18) gives us the equilibrium distribution of plant productivities. From this productivity density function and the assumption of $a' > 0$, it is clear that the increase in the number of operating plants slides the distribution rightward while maintaining its form.

**Implication 3** (*Agglomeration effects*)**.** An increase in the number of operating plants in a region $N_p$ slides the productivity distribution to the right.

### 3.4 Empirical hypotheses and identification strategy

Based on the theoretical model and its three implications, we can consider the following four cases with respect to the channels of productivity improvement. For expositional simplicity, we consider two regions $r = c$ (cluster) and $r = n$ (non-cluster).

**Case 1** (*Only the selection effect matters*)**.** When there is no agglomeration effect, only selection affects productivity. In this case, $a(N_p) = 1$ holds for any value of $N_p$. On the other hand, selection implies that $\hat{h}_c < \hat{h}_n$, where $\hat{h}_c$ ($\hat{h}_n$) is the cutoff unit labor requirement in region $c(n)$. This raises the log productivity cut-off in the cluster: $\hat{\phi}_c > \hat{\phi}_n$. This case is represented in Figure 2(a). The solid line represents the log productivity distribution in the cluster, while the dashed line represents that in the non-cluster. The log productivity distribution in the cluster is left truncated.

**Case 2** (*Only the agglomeration effect matters*)**.** In this case, only the agglomeration effect improves the plants' productivity. To eliminate the selection effects, we impose $s_c = s_n = s$, where $s_c$ ($s_n$) is the startup cost in region $c$ ($n$). Then, the intention of the selection is the same in the both cluster and non-cluster, and therefore, $\hat{h}_c = \hat{h}_n$ and $N_{pc} = N_{pn}$, where $N_{pc}$ ($N_{pn}$) is the number of plants in region $c$ ($n$). In order to establish clusters and non-clusters,

[3]Indeed, the performance-based wage system was prevalent in the silk-reeling clusters (Nakabayashi, 2003).

we need to assume $N_c > N_n$ by exogenous reasons that are outside the scope of our model. Only firms in clusters benefit from larger worker interactions, $\ln[a(N_{pc})] > \ln[a(N_{pn})]$. Thus, the log productivity simply slides to the right while maintaining its distribution form. This case is shown in Figure 2(b).

**Case 3** (*Both the selection and agglomeration effects matter*)**.** In this case, the fixed entry costs are different between the cluster and non-cluster, $s_c < s_n$, and the concentration of workers improves their productivity, $a' > 0$ and $a'' < 0$. Thus, both left truncations by selection and right slide by agglomeration occur in the cluster. This case is shown in Figure 2(c).

**Case 4** (*Neither effect matters*)**.** In this case, the fixed entry costs are the same for all regions and the concentration of workers does not improve their productivity, $a(N_p) = 1$. Then, log productivity distribution is common across regions. Thus, there is no difference in the productivities across regions. This case is shown in Figure 2(d).

= Figures 2(a) to 2(d) =

To distinguish these four cases empirically, we employ two measures that characterize the productivity distributions. The first measure is the interquartile range of the distribution. If no selection effect exists (cases 2 and 4), the shape of the distribution should be the same for clusters and non-clusters and therefore, the interquartile range should have no difference. On the other hand, if a selection effect exists (cases 1 and 3), the productivity distribution should be left truncated in clusters and the interquartile range should be smaller than for non-clusters. Hence, by comparing the interquartile range between clusters and non-clusters, we would be able to detect the presence of the selection effect.[4]

The second measure consists of the percentiles of the distribution. Since selection left truncates the distribution, we should observe a rise of lower percentile points rather than higher percentile points of log-productivity distribution. On the other hand, the agglomeration effect affects every percentile point of the distribution because it shifts the whole distribution rightwards. Thus, if the agglomeration effects are in place, both the higher and lower percentile points of the distribution should increase.

The above discussions are summarized in Table 1.

= Table 1 =

The table represents the direction of the shift of each measure of distribution in clusters relative to non-clusters for four cases.

# 4 Identification of agglomeration and selection with productivity distribution

Based on the theoretical prediction summarized in Table 1, this section identifies the agglomeration and selection effects by focusing on the characteristics and shift of productivity distributions.

[4]Of course, variance is also an informative measure of truncation. However, empirically, the interquartile range is more robust for outliers.

### 4.1 Data

We compiled the data of the silk-reeling industry census data, *Zenkoku Seishi Kojo Chosa*, for the two data points 1908 and 1915. The data include plant-level information of plant name, location, year of foundation, number of pots, number of workers, number of business days per year, type of powers, and output. This data set covers both hand-reeling and machine-reeling plants. Here, we focus on machine reeling plants because hand-reeling and machine-reeling are completely different techniques. We matched plants over two periods and constructed an unbalanced panel data set by using the plant name, location, and year of foundation. The number of machine-reeling plants in 1908 was 2385 and that in 1915 was 2263. The number of plants that existed in 1908 and survived until 1915 was 910. The extent of agglomeration is measured by regional plant density, computed as number of plants per km$^2$. Information regarding the regional area at the prefecture or county level was obtained from the GIS (Geographical Information System) data for 1937, from "Taisho-Showa Gyoseikai Data."[5]

### 4.2 Measures of plant-level productivity

As discussed in the previous section, we focus on the shape of the productivity distributions to distinguish the channels of productivity improvement effects. For this purpose, we first estimate the productivity of each plant. We use Total Factor Productivity (TFP) as the primary measure of plant-level productivity. In order to estimate TFP, we specify the firm-level production function as below:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \delta Z_{it} + \omega_{it} + \epsilon_{it}, \tag{19}$$

where $y_{it}$ is the log of output, $k_{it}$ is the log of capital input (number of pots), $l_{it}$ is the log of labor inputs (number of female workers), $m_{it}$ is the log of intermediate input (quantity of cocoons used), $Z_{it}$ is the vector of the plant $i$'s observable characteristics, $\omega_{it}$ and $\epsilon_i$ are the productivity terms that are unobservable to the econometrician. While $\epsilon_{it}$ is also unobserved by firms before they make their input decisions, $\omega_{it}$ is observable. We assume $\omega_{it} = \omega_i$, that is, the observable productivity for plants does not change through the study period (7 years). As the plant $i$'s observable characteristics, we include the log of plant age, a dummy variable indicating that the plant used water power, and a dummy variable indicating that the plant used steam power. Given this assumption, we estimate this production function by the fixed effect model. For a robustness check, we report the results using TFP estimated with the method proposed by Levinsohn and Petrin (2003) that relaxes the assumption of $\omega_{it} = \omega_i$ in Appendix A. Our main results are unchanged with the application of this alternative method.

The estimation results are reported in Table 2.

= Table 2 =

The coefficients of ln (capital) and ln (worker) have expected signs and magnitudes with high statistical significance. The coefficients of ln (intermediate) is positive, but not significantly different from zero. We interpret $y_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_l l_{it} + \hat{\beta_m} m_{it} - \hat{\delta} Z_{it}$ as the TFP of plant $i$ in period $t$. We also use the output per pot (capital productivity) and output per worker (labor productivity) as alternative measures of plant-level productivity. In the Japanese silk-reeling industry, the output per pot and output per worker were conventionally used as measures to evaluate plant-level productivity.

[5]The data is published by Yuji Murayama, Division of Spatial Information Science, Graduate School of Life and Environmental Sciences, University of Tsukuba, URL: http://giswin.geo.tsukuba.ac.jp/teacher/murayama/data_map.html

### 4.3 Main result

We now proceed to distinguish between agglomeration effects and selection effects by examining the shape of the plant-level productivities estimated in the previous section with the theoretical predictions in section 3.4.

We use a prefecture as a unit of observation of regional productivity density to obtain sufficient observations. All prefectures in Japan are classified into two groups on the basis of regional plant densities (number of plants per $km^2$ in a region). Prefectures with plant densities higher than the median value are classified into the *clustered prefectures*, and the other prefectures are classified into the *non-clustered prefectures*.

First, we estimate the kernel density functions[6] of the plant-level productivity for each group of prefectures. Figure 3 represents the kernel densities.

= Figure 3 =

The solid line refers to the density of the clustered prefectures and the dashed line refers to the density of the non-clustered prefectures. In every figure, the estimated density in the lower tail of the distribution is lower for the clustered prefectures than for the non-clustered prefectures, while the density in the higher tail of the distributions is similar for the two prefecture groups. Moreover, the shapes and positions of the two distributions seem to be similar except for the lower tails. There is no clear sign that the distribution of the clustered prefectures slides to the right. Given the predictions summarized in Table 1, these features of the two distributions suggest that a selection effect existed but an agglomeration effect did not.

These observations are also confirmed by the descriptive statistics of the productivity distributions shown in Table 3.

= Table 3 =

Panel A represents the statistics for 1908 and Panel B represents those for 1915. The mean productivities are significantly higher in the cluster for every productivity measure and period, indicating that plants located in the industrial clusters were indeed more productive than those in the non-clusters.

Our theoretical predictions in section 3.4 and Table 1 suggest that the interquartile range is informative in detecting selection effects. If the productivity distribution is truncated by selection, we should observe a shorter interquartile range. We can also examine the existence of selection and agglomeration effects by looking at the percentiles of distribution at lower and higher tails: while selection affects only the lower tail, agglomeration affects every support of the distribution because agglomeration shifts the whole distribution rightwards. Hence, if the agglomeration effect exists, both lower and higher tails of the distribution should shift to the right.

Table 3 reveals that the interquartile range of the distribution in the clustered prefectures is smaller than that in the non-clustered prefectures. This suggests the truncation of the distribution implied by plant-selection. This interpretation is further supported by the percentiles of the distribution. In the lower tails (the 10th and 25th percentiles), percentiles in clusters are much higher than in non-clusters. For example, in 1908, while the 10th percentile in the cluster is –0.21, the same percentile in the non-cluster is –0.33, and the difference is 0.1. On the other

[6]We estimate the density functions by using the Epanechnikov kernel with optimal bandwidth.

hand, in higher tails, percentiles in both the cluster and non-cluster are quite similar (the 90th percentile in both cluster and non-cluster are 0.18). These findings are consistent with Case 1 in Table 1: productivity distribution is left truncated but no right shift is observed. According to our theoretical prediction, this implies the existence of plant-selection and non-existence of the agglomeration effect.

Next, we econometrically distinguish between the agglomeration and selection effects by utilizing prefectural variations. We first investigate the effect of plant-densities on the interquartile range of productivity distribution. We index each prefecture by $p$, and estimate the following equation:

$$\mathrm{IQR}_{pt} = \alpha + \beta \ln(\mathrm{D}_{pt}) + \mathrm{year}_t + \varepsilon_{pt} \tag{20}$$

where $\mathrm{IQR}_{pt}$ refers to the interquartile range of plant-level productivity distribution in prefecture $p$ in period $t$, $\mathrm{D}_{pt}$ is the plant density and $\mathrm{year}_t$ is the year fixed effects. Under the presence of selection effects, an increase in the plant density will truncate the distribution and shorten the interquartile range; thus, we expect a negative sign for $\beta$. We estimate this equation by pooled OLS with year fixed effects. Because this estimation focuses on the productivity distribution and requires a certain number of observations (plants) in each prefecture, we restrict samples to prefectures that had more than 20 plants. The results are shown in columns (1) to (3) in Table 4.

= Table 4 =

In every result (columns 1–3), the coefficients of plant density are negative and statistically significant. This suggests the truncation of productivity distributions, which is consistent with the existence of the selection effect.

Next, we examine the role of the agglomeration effect by focusing on the percentiles of productivity distribution. We estimate the following equation:

$$\mathrm{P}^u_{pt} = \alpha + \beta \ln(\mathrm{D}_{pt}) + \mathrm{year}_t + \varepsilon_{pt}, \tag{21}$$

where $\mathrm{P}^u_{pt}$ is the $u$-th percentile of the log productivity distribution in prefecture $p$ for period $t$. The equation is estimated by pooled OLS.

The results are shown in Table 5.

= Table 5 =

Columns (1) to (4) use TFP as a measure of productivity while columns (5) to (8) and (9) to (12) use output per pot and output per worker respectively. Regardless of the measure of productivity, coefficients of plant density are positive and significant for the lower tail, that is, the 10th and 25th percentiles (columns 1, 2, 5, 6, 9, and 10). This result is consistent with both the agglomeration effect and selection effects. On the other hand, every coefficient of plant density is not statistically different from zero at the 75th or 90th percentiles (columns 3, 4, 7, 8, 11, and 12). This implies that higher plant density had no effect on the productivity distribution shifting rightwards. The evidence runs contrary to the existence of the agglomeration effect.

These results indicate that the increase of plant density truncated the log productivity distribution in the lower tail but had no effect in shifting the distribution rightwards. This is consistent with the existence of the selection effect and non-existence of the agglomeration effect (Case 1 in Table 1).

# 5 Learning and productivity growth in clusters

While the analyses on the shape of productivity distribution provide a number of interesting findings, their validity in distinguishing between selection and agglomeration effects rests on the assumption of how agglomeration affects productivity. One important assumption is that agglomeration affects every plant in a cluster, and hence, it slides the distribution to the right but never skews the distribution form. However, plant agglomeration might also skew the distribution. For example, consider one important component of the agglomeration effect: learning. Low-productivity plants might learn and improve their productivities faster than high-productivity plants. Such catch-up will skew the productivity distribution in the same manner as the selection effect. If low-productivity plants in clusters can learn faster from many surrounding plants than those in non-clusters, left truncation does not necessarily imply the existence of selection.

Therefore, we take a different approach to distinguish between the agglomeration and selection effects by focusing on the timing when productivity growth occurred. If the agglomeration effect is in place and learning from leading plants improves plant-level productivities, we should observe faster productivity growth in clusters than in non-clusters *after* the plants are operational. On the other hand, if truncation is caused by selection *before* operation, then the productivity distribution in clusters should be truncated even for young plants and they should be already more productive on an average from the start-up.

## 5.1 Productivity growth after operation

We first examine the first hypothesis on the productivity growth after operation by comparing the productivity growth rates between plants in clusters and non-clusters.

Descriptive statistics of productivity growth rate from 1908 to 1915 for three different productivity measures are shown in Table 6.

= Table 6 =

Younger plants might tend to learn and improve their productivity faster than older plants. We report descriptive statistics for start-up plants separately. We define *start-up* plants as plants with age less than five years.

For every productivity measure, the average productivity growth rate is similar for the cluster and non-cluster. Rather, the average growth rate in the cluster is smaller than in the non-cluster. We test the difference of average productivity growth rate between the cluster and non-cluster using a $t$-test, but the null hypothesis that the mean difference of growth rate between the cluster and non-cluster is zero is not rejected at the conventional levels of statistics for every measure of productivity. Furthermore, this relationship also holds even if we restrict samples to *start-up* plants. Overall, start-up plants grew faster than older plants in terms of every productivity measure. Meanwhile, the average productivity growth rate of *start-up* plants in the cluster is again lower than in the non-cluster, and we also cannot reject the null hypothesis that the mean difference of the growth rate between the cluster and non-cluster is zero in every measure of productivity. Thus, there is no evidence that plants learned faster in clusters than in non-clusters.

To control plant-level differences, we econometrically test the learning effects in clusters, by estimating the following equation,

$$\text{GrowthRate}_{icp} = \alpha + \beta D_{cp} + \delta Z_{icp} + \gamma \text{Productivity1908}_{icp} + \text{pref}_p + \varepsilon_{icp}, \tag{22}$$

where $\text{GrowthRate}_{icp}$ is the productivity growth rate of plant $i$ located in county $c$ in prefecture $p$ from 1908 to 1915, $D_{cp}$ is the plant density at county level, and $Z_{icp}$ is the vector of plant-level control variables (number of pots, number of workers, age, steam power dummy, water power dummy). To control low-productivity plants' faster learning and growth (catch-up), we include the initial productivity in 1908. Table 7 reports the results.

= Table 7 =

Again, we use three measures for the plant-level productivity. Column (1) is the baseline result. Even after controlling plant-level variables, the coefficient of plant density is not statistically different from zero. That is, plants in cluster did not improve their productivity faster than plants in non-clusters. Column (2) controls initial productivity in 1908. Interestingly, the coefficient of initial productivity in 1908 is negative and significant. This suggests a catch-up by low-productivity plants. However, the coefficient of plant density is still not statistically different from zero. Low-productivity plants did learn and catch-up, but their speed was not significantly different for clusters and non-clusters. This result is robust to the alternative measures of productivity (columns 3–6) or restricted samples of *start-up* plants (columns 7–12). The coefficient of initial productivity in 1908 is significantly negative but the plant density is not statistically different from zero. In sum, these results are against the presence of learning implied by the agglomeration effect. Low-productivity plants did catch-up, but we find no evidence that plants in clusters improved their productivity faster than their counterparts in non-clusters.

## 5.2 Productivity distribution of start-up plants

We now proceed to examine the second hypothesis that the productivity distribution of younger plants was already truncated and they were more productive on an average if selection took place before operation.

Table 8 reports the descriptive statistics of plant-level productivities for start-up plants (plants with age less than five years).

= Table 8 =

We report the statistics for 1908 and 1915 in panels A and B respectively. For every measure of productivity, the average productivity of start-up plants is higher in clusters than in non-clusters for both periods. Moreover, for every measure and period, the interquartile range is narrower for clusters than for non-clusters, except for output per labor in 1915. This implies that productivity distribution is more severely truncated in clusters even for restricted samples of start-up plants. Furthermore, lower percentiles are much higher for clusters, while higher percentiles are similar between clusters and non-clusters implying that there was no shift of distribution. Overall, even for the restricted samples of new start-up plants, the distribution in clusters was already left truncated. This suggests that the left truncation of the distribution occurred before the plants were operational, which is consistent with the selection under the assumption of rational agents broadly assumed in the models of selection (Melitz, 2003; Syverson, 2004; Melitz and Ottaviano, 2008).

## 6 Concluding remarks

In this paper, we attempted to distinguish the two channels through which industrial clusters improved plant-level productivity, focusing on the Japanese silk-reeling industry in the period from 1908 to 1915. On the basis of nested model of selection and agglomeration, we considered the agglomeration effect, which improves the productivities of all the plants in a region and the plant-selection effect, which raised the average regional productivity by expelling less productive plants through intense competition.

Using plant-level data, we distinguished the channels of productivity improvement on the basis of the theoretical predictions on the characteristics of productivity distribution: the selection effect would left truncate the productivity distribution, while the agglomeration effect slides the distribution rightwards without changing its form. We found that the interquartile range of the productivity distribution was significantly smaller for the clustered prefectures than for the non-clustered prefectures, but found no evidence of a rightward slide of the distribution for the clustered prefectures. This suggests the existence of the selection effect and the non-existence of the agglomeration effect. The result is further supported by the findings that there was no difference in the productivity growth rate of individual plants between clusters and non-clusters and the productivity distribution of start-up plants in clusters was already truncated. These two results suggest that productivity growth through learning suggested by the agglomeration effect was not evident, while the truncation of less productive plants in clusters occurred before the operation. In this sense, truncation and average productivity improvement effects in clusters principally occurred through the selection process, which expelled the less productive plants in clusters before they started-up.

We therefore conclude that in the Japanese silk-reeling industry, higher average productivity in clusters was not caused by the agglomeration effect but through selection, that is, the intensification of competition in clusters expelled the low-productivity plants, and consequently, only relatively more productive plants survived. This suggests the importance of competition in improving productivity in specialized industrial clusters.

## References


Baldwin, R., T. Okubo (2006), Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting, *Journal of Economic Geography* 6, 323–346

Behrens, K., G. Mion, Y. Murata, and J. Südekum (2009), Trade, Wages, and Productivity, *CEPR Discussion Paper Series*, #7369.

Ciccone, A. (2002), Agglomeration Effects in Europe, *European Economic Review* 46, 213–227.

Ciccone, A., R. Hall (1996), Productivity and the Density of Economic Activity, *American Economic Review* 86, 54–70.

Combes, P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2009), The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection, *mimeo.*

Corcos, G., M. Del Gatto, G. Mion, and G. Ottaviano (2010), Productivity and Firm Selection: Quantifying the "New" Gains from Trade, *mimeo.*

Duran, L. (1913) *Raw silk: A Practical Hand-Book for the Buyer.* Silk Publishing Company, New York.

Fujita, M., H. Ogawa (1982), Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations, *Regional Science and Urban Economics* 12, 161–196.

Glaeser, E., H. Kallal, J. Shainkman, and A. Shleifer (1992), Growth in Cities, *Journal of Political Economy* 100, 1126–1152.

Henderson, V. (2003), Marshall's Scale Economies, *Journal of Urban Economics* 53, 1–28.

Hirano, Y. (1990) *Kindai yo-san-gyo no hatten to kumiai seishi (Development of modern sericulture and cooperative silk-reeling).* University of Tokyo Press, Tokyo. (in Japanese).

Ishii, K. (1972) *Nihon sanshigyo-shi bunseki (Analysis of Japanese Silk industry).* University of Tokyo Press, Tokyo. in Japanese.

Jacobs, J. (1969) *The Economy of Cities.* Vintage, New York.

Kambayashi, R. (2001) "Tokyu chingin seido to ko-jo toroku seido," (Grade wage system and registration system of female factory workers) in T. Okazaki ed., Torihiki seido no keizai-shi (Governing business transactions: A historical perspective), University of Tokyo Press, Tokyo. (in Japanese).

Levinsohn, J., A. Petrin (2003), Estimating Production Functions Using Inputs to Control for Unobservables, *The Review of Economic Studies* 70, 317–341.

Lucas, R., E. Rossi-Hansberg (2002), On the Internal Structure of Cities, *Econometrica* 70, 1445–1476.

Matsumura, S. (1992) *Senkan-ki nihon sannshigyo-shi kenkyu (Historical study of Japanese silk-reeling in the inter-war period).* University of Tokyo Press, Tokyo. (in Japanese).

Melitz, M. (2003), The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica* 71, 1695–1725.

Melitz, M., G. Ottaviano (2008), Market Size, Trade, and Productivity, *Review of Economic Studies* 75, 295–316.

Melo, P., D. Graham, and R. Noland (2009), A Meta-Analysis of Estimates of Urban Agglomeration Economies, *Regional Science and Urban Economics* 39, 332–342.

Nakabayashi, M.(2003), *Kindai shihonshugi no soshiki: Seishigyo no hatten ni okeru torihiki no tochi to seisan no kozo (An orgazation of modern capitalism: The governance of trade and the system of production in the development of the silk reeling industry).* University of Tokyo Press, Tokyo. (in Japanese).

Rosenthal, S., W. Strange (2004), Evidence on the Nature and Sources of Agglomeration Economies, in V. Henderson and J. Thisse (eds.) *Handbook of Regional and Urban Economics*, Vol. 4. North-Holland, Amsterdam, 2119–2171.

Syverson, C. (2004), Market Structure and Productivity: A Concrete Example, *Journal of Political Economy* 112, 1181–1222.

Tojo, Y.(1990), *Seishi domei no joko toroku seido (Worker registration system of silk-reeling association).* University of Tokyo Press, Tokyo. (in Japanese).

Table 1: Measures of distribution in clusters relative to non-clusters

| Cases | Mean | Interquartile range | Lower percentile | Higher percentile |
|---|---|---|---|---|
| Case 1: Selection | + | − | + | + |
| Case 2: Agglomeration | + | 0 | + | ++ |
| Case 3: Selection & agglomeration | + | − | + | ++ |
| Case 4: Neither effect | 0 | 0 | 0 | 0 |

Table 2: Estimation result of plant-level productivity

| ln(capital) | ln(labor) | ln(intermediates) | ln(age) | Water power dummy | Steam power dummy | Constant | No. obs. | $R^2$ |
|---|---|---|---|---|---|---|---|---|
| 0.152** | 0.153** | 0.719** | 0.056** | 0.053* | 0.004 | 2.170 ** | 4479 | 0.955 |
| (0.056) | (0.074) | (0.076) | (0.022) | (0.032) | (0.035) | (0.134) | | |

Note: The dependent variable represents the log of output. The capital represents the number of pots, labor represents the number of workers, and intermediates represents the quantity of cocoons. Estimated by OLS with plant fixed effects.
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 3: Descriptive statistics of plant-level productivity

| Panel A: 1909 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP | | | | | | | | |
| | All | -0.03 | 0.22 | -0.27 | -0.14 | -0.01 | 0.09 | 0.18 |
| | Cluster | -0.002 | 0.19 | -0.21 | -0.09 | 0.01 | 0.1 | 0.18 |
| | Non-cluster | -0.05 | 0.26 | -0.33 | -0.18 | -0.04 | 0.08 | 0.18 |
| | | (4.288) | | | | | | |
| Output per pot | | | | | | | | |
| | All | 3.95 | 0.71 | 3.2 | 3.62 | 4.03 | 4.33 | 4.57 |
| | Cluster | 4.07 | 0.57 | 3.44 | 3.82 | 4.07 | 4.39 | 4.57 |
| | Non-cluster | 3.82 | 0.82 | 2.96 | 3.43 | 3.92 | 4.25 | 4.56 |
| | | (10.031) | | | | | | |
| Output per Worker | | | | | | | | |
| | All | 3.92 | 0.66 | 3.16 | 3.63 | 4.00 | 4.29 | 4.51 |
| | Cluster | 4.04 | 0.55 | 3.44 | 3.81 | 4.06 | 4.35 | 4.54 |
| | Non-cluster | 3.8 | 0.79 | 2.96 | 3.44 | 3.93 | 4.23 | 4.47 |
| | | (10.113) | | | | | | |
| Panel B: 1916 | | | | | | | | |
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP | | | | | | | | |
| | All | 0.03 | 0.28 | -0.27 | -0.1 | 0.06 | 0.18 | 0.29 |
| | Cluster | 0.04 | 0.24 | -0.2 | -0.07 | 0.07 | 0.18 | 0.28 |
| | Non-cluster | 0.004 | 0.34 | -0.34 | -0.15 | 0.05 | 0.19 | 0.29 |
| | | (3.218) | | | | | | |
| Output per pot | | | | | | | | |
| | All | 4.17 | 0.87 | 3.22 | 3.78 | 4.29 | 4.65 | 4.93 |
| | Cluster | 4.26 | 0.79 | 3.53 | 3.87 | 4.34 | 4.66 | 4.93 |
| | Non-cluster | 4.05 | 1.04 | 2.93 | 3.59 | 4.19 | 4.63 | 4.94 |
| | | (7.180) | | | | | | |
| Output per Worker | | | | | | | | |
| | All | 4.14 | 0.83 | 3.24 | 3.76 | 4.26 | 4.6 | 4.86 |
| | Cluster | 4.23 | 0.77 | 3.49 | 3.85 | 4.31 | 4.61 | 4.88 |
| | Non-cluster | 4.02 | 0.94 | 2.91 | 3.62 | 4.18 | 4.56 | 4.83 |
| | | (7.581) | | | | | | |

Note: TFP is estimated by the fixed effect OLS. The $t$-values of the $t$-test are presented in parentheses. The null hypotheses of the $t$-test is that the average productivities is not different for clusters and non-clusters.

Table 4: Plant density and interquartile range of productivity distribution

| | (1) | (2) | (3) |
|---|---|---|---|
| Measure of Productivity | TFP | Output per pot | Output per worker |
| ln(density) | -0.0276* | -0.129** | -0.130** |
| | (0.0151) | (0.0453) | (0.0539) |
| Constant | 0.256** | 0.860** | 0.811** |
| | (0.0280) | (0.0876) | (0.104) |
| Observations | 45 | 45 | 45 |
| Adjusted $R^2$ | 0.071 | 0.111 | 0.102 |

Note: The dependent variables are the interquartile range of the productivity distribution at the prefectural level.
TFP is estimated by the fixed effect OLS.
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 5: Plant density and percentiles of productivity distribution

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure of Productivity Percentiles | TFP 10th | TFP 25th | TFP 75th | TFP 90th | Output per Pot 10th | Output per Pot 25th | Output per Pot 75th | Output per Pot 90th | Output per Worker 10th | Output per Worker 25th | Output per Worker 75th | Output per Worker 90th |
| ln(density) | 0.0506* | 0.0430* | 0.0154 | 0.0189 | 0.197** | 0.181** | 0.0527 | 0.0546 | 0.191** | 0.169** | 0.0387 | 0.0672 |
| | (0.0290) | (0.0243) | (0.0155) | (0.0141) | (0.0772) | (0.0764) | (0.0560) | (0.0475) | (0.0801) | (0.0804) | (0.0535) | (0.0473) |
| Constant | -0.286** | -0.168** | 0.0875** | 0.166** | 3.122** | 3.432** | 4.292** | 4.557** | 3.138** | 3.453** | 4.264** | 4.473** |
| | (0.0509) | (0.0422) | (0.0254) | (0.0224) | (0.139) | (0.136) | (0.0901) | (0.0788) | (0.147) | (0.145) | (0.0825) | (0.0724) |
| Year fixed effect | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |
| Adjusted $R^2$ | 0.031 | 0.046 | 0.145 | 0.292 | 0.115 | 0.117 | 0.202 | 0.318 | 0.075 | 0.087 | 0.227 | 0.385 |

Note: Dependent variables are percentile points of the productivity distribution at the prefectural level. TFP is estimated by the fixed effect OLS.
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 6: Plant-level productivity growth rate, 1909 to 1916

| Panel A. Measure of productivity: TFP | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | All plants | | | | Start-up plants | | | |
| Cluster vs. Non-cluster | Obs. | Mean | Standard deviation | t-value | Obs. | Mean | Standard deviation | t-value |
| Cluster | 448 | 0.279 | 0.680 | | 115 | 0.417 | 0.829 | |
| Non-cluster | 461 | 0.311 | 1.053 | | 83 | 0.640 | 1.842 | |
| All | 909 | 0.295 | 0.888 | | 198 | 0.510 | 1.350 | |
| Difference (Cluster/Non-cluster) | | 0.032 | 0.059 | 0.535 | | 0.223 | 0.194 | 1.148 |
| **Panel B. Measure of productivity: Output per pot** | | | | | | | | |
| | All plants | | | | Start-up plants | | | |
| Cluster vs. Non-cluster | Obs. | Mean | Standard deviation | t-value | Obs. | Mean | Standard deviation | t-value |
| Cluster | 448 | 0.471 | 1.011 | | 115 | 0.676 | 1.219 | |
| Non-cluster | 461 | 0.494 | 1.568 | | 83 | 1.053 | 3.014 | |
| All | 909 | 0.482 | 1.323 | | 198 | 0.834 | 2.163 | |
| Difference (Cluster/Non-cluster) | | 0.023 | 0.088 | 0.265 | | 0.377 | 0.311 | 1.213 |
| **Panel C. Measure of productivity: Output per worker** | | | | | | | | |
| | All plants | | | | Start-up plants | | | |
| Cluster vs. Non-cluster | Obs. | Mean | Standard deviation | t-value | Obs. | Mean | Standard deviation | t-value |
| Cluster | 448 | 0.459 | 1.049 | | 115 | 0.652 | 1.198 | |
| Non-cluster | 461 | 0.489 | 1.551 | | 83 | 1.009 | 2.973 | |
| All | 909 | 0.474 | 1.327 | | 198 | 0.802 | 2.131 | |
| Difference (Cluster/Non-cluster) | | 0.030 | 0.088 | 0.337 | | 0.357 | 0.307 | 1.164 |

Note: TFP is estimated by the fixed effect OLS. t-value columns show the results of the *t*-test. The null hypotheses of the *t*-test is that the average productivity growth is not different for clusters and non-clusters.

Table 7: Estimation of growth effects

| Dependent: Measure of productivity growth | (1) TFP | (2) TFP | (3) Output per pot | (4) Output per pot | (5) Output per worker | (6) Output per worker | (7) TFP | (8) TFP | (9) Output per pot | (10) Output per pot | (11) Output per worker | (12) Output per worker |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Samples | All | All | All | All | All | All | Start-up | Start-up | Start-up | Start-up | Start-up | Start-up |
| ln(density) | -0.0148 | -0.00968 | -0.0138 | -0.00820 | 0.0105 | 0.0165 | -0.0589 | -0.0911 | -0.119 | -0.165 | -0.0949 | -0.142 |
| | (0.0284) | (0.0258) | (0.0410) | (0.0389) | (0.0391) | (0.0370) | (0.0919) | (0.0832) | (0.151) | (0.145) | (0.147) | (0.143) |
| ln(pot) | 0.968** | 0.877** | 1.435** | 0.300 | -0.691** | -0.600** | 2.393** | 2.240** | 3.552** | 1.989** | 0.322 | 0.346 |
| | (0.393) | (0.377) | (0.585) | (0.634) | (0.259) | (0.229) | (0.438) | (0.432) | (0.605) | (0.950) | (0.436) | (0.439) |
| ln(worker) | -0.340 | -0.274 | -0.391 | -0.130 | 1.751** | 0.732** | -0.816 | -0.806 | -0.914 | -0.657 | 2.063* | 0.700 |
| | (0.418) | (0.389) | (0.652) | (0.644) | (0.466) | (0.318) | (0.638) | (0.573) | (1.129) | (1.101) | (1.076) | (0.924) |
| ln(cocoon) | -0.552** | -0.511** | -0.962** | -0.0380 | -1.009** | -0.0282 | -1.472** | -1.297** | -2.535** | -1.144 | -2.407** | -0.981 |
| | (0.183) | (0.134) | (0.309) | (0.200) | (0.317) | (0.194) | (0.554) | (0.479) | (1.113) | (0.953) | (1.114) | (0.912) |
| ln(age) | -0.0783* | -0.133** | -0.128* | -0.121* | -0.132* | -0.124* | -0.182 | -0.231 | -0.437 | -0.417 | -0.498 | -0.477 |
| | (0.0415) | (0.0431) | (0.0714) | (0.0691) | (0.0721) | (0.0697) | (0.222) | (0.206) | (0.400) | (0.369) | (0.398) | (0.368) |
| Steam power dummy | 0.0531 | 0.0933 | 0.0720 | 0.122 | 0.101 | 0.154 | 0.398 | 0.370 | 0.542 | 0.508 | 0.688* | 0.653 |
| | (0.0791) | (0.0753) | (0.114) | (0.111) | (0.120) | (0.115) | (0.262) | (0.257) | (0.418) | (0.413) | (0.407) | (0.403) |
| Water power dummy | -0.0316 | -0.0732 | -0.0214 | -0.00273 | 0.00137 | 0.0212 | 0.701* | 0.671** | 1.395** | 1.437** | 1.421** | 1.465** |
| | (0.0811) | (0.0764) | (0.128) | (0.127) | (0.131) | (0.129) | (0.355) | (0.339) | (0.692) | (0.687) | (0.692) | (0.687) |
| ln(TFP) in 1909 | | -1.100** | | | | | | -1.114** | | | | |
| | | (0.203) | | | | | | (0.483) | | | | |
| ln(Output per pot) in 1909 | | | | -1.221** | | | | | | -1.587* | | |
| | | | | (0.273) | | | | | | (0.822) | | |
| ln(Output per worker) in 1909 | | | | | | -1.296** | | | | | | -1.627** |
| | | | | | | (0.277) | | | | | | (0.780) |
| Constant | 1.370** | 1.337** | 2.567** | 5.179** | 2.781** | 5.554** | 2.870** | 2.401** | 5.301** | 8.077** | 5.433** | 8.279** |
| | (0.420) | (0.350) | (0.737) | (1.142) | (0.757) | (1.175) | (1.225) | (1.114) | (2.489) | (3.413) | (2.499) | (3.382) |
| Observations | 909 | 909 | 909 | 909 | 909 | 909 | 198 | 198 | 198 | 198 | 198 | 198 |
| Adjusted $R^2$ | 0.157 | 0.265 | 0.160 | 0.219 | 0.159 | 0.226 | 0.534 | 0.574 | 0.405 | 0.435 | 0.401 | 0.435 |

Note: The dependent variables represent the growth rate of the plant-level productivity from 1909 to 1916. TFP is estimated by the fixed effect OLS. A start-up is a plant with age less than five years.
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 8: Descriptive statistics of start-up plants' productivities

| Panel A: 1908 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP | | | | | | | | |
| | All | 0.022 | 0.218 | -0.232 | -0.084 | 0.037 | 0.134 | 0.207 |
| | Cluster | 0.048 | 0.182 | -0.168 | -0.041 | 0.058 | 0.141 | 0.206 |
| | Non-cluster | -0.008 | 0.265 | -0.269 | -0.141 | 0.014 | 0.124 | 0.214 |
| | | (2.553) | | | | | | |
| Output per pot | | | | | | | | |
| | All | 3.873 | 0.670 | 3.037 | 3.560 | 3.974 | 4.230 | 4.491 |
| | Cluster | 3.967 | 0.526 | 3.337 | 3.753 | 4.030 | 4.279 | 4.489 |
| | Non-cluster | 3.765 | 0.839 | 2.931 | 3.338 | 3.861 | 4.176 | 4.510 |
| | | (4.318) | | | | | | |
| Output per worker | | | | | | | | |
| | All | 3.855 | 0.655 | 3.050 | 3.567 | 3.976 | 4.222 | 4.435 |
| | Cluster | 3.955 | 0.496 | 3.306 | 3.757 | 4.014 | 4.253 | 4.463 |
| | Non-cluster | 3.740 | 0.871 | 2.885 | 3.302 | 3.883 | 4.173 | 4.402 |
| | | (4.626) | | | | | | |

| Panel A: 1915 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP | | | | | | | | |
| | All | 0.065 | 0.313 | -0.232 | -0.078 | 0.099 | 0.235 | 0.346 |
| | Cluster | 0.072 | 0.296 | -0.206 | -0.067 | 0.100 | 0.229 | 0.342 |
| | Non-cluster | 0.054 | 0.342 | -0.268 | -0.101 | 0.098 | 0.241 | 0.349 |
| | | (0.823) | | | | | | |
| Output per pot | | | | | | | | |
| | All | 4.088 | 0.931 | 3.211 | 3.650 | 4.135 | 4.581 | 4.889 |
| | Cluster | 4.122 | 0.932 | 3.361 | 3.662 | 4.168 | 4.594 | 4.871 |
| | Non-cluster | 4.034 | 0.951 | 3.056 | 3.622 | 4.097 | 4.573 | 4.932 |
| | | (1.718) | | | | | | |
| Output per worker | | | | | | | | |
| | All | 4.056 | 0.890 | 3.211 | 3.636 | 4.104 | 4.525 | 4.850 |
| | Cluster | 4.091 | 0.898 | 3.373 | 3.628 | 4.117 | 4.525 | 4.885 |
| | Non-cluster | 3.998 | 0.870 | 3.037 | 3.644 | 4.092 | 4.514 | 4.828 |
| | | (1.859) | | | | | | |

Note: The $t$-values of the $t$-test are in parentheses. The null hypotheses of the $t$-test is that the average productivities are not different for clusters and non-clusters. A start-up is a plant with age less than five years.

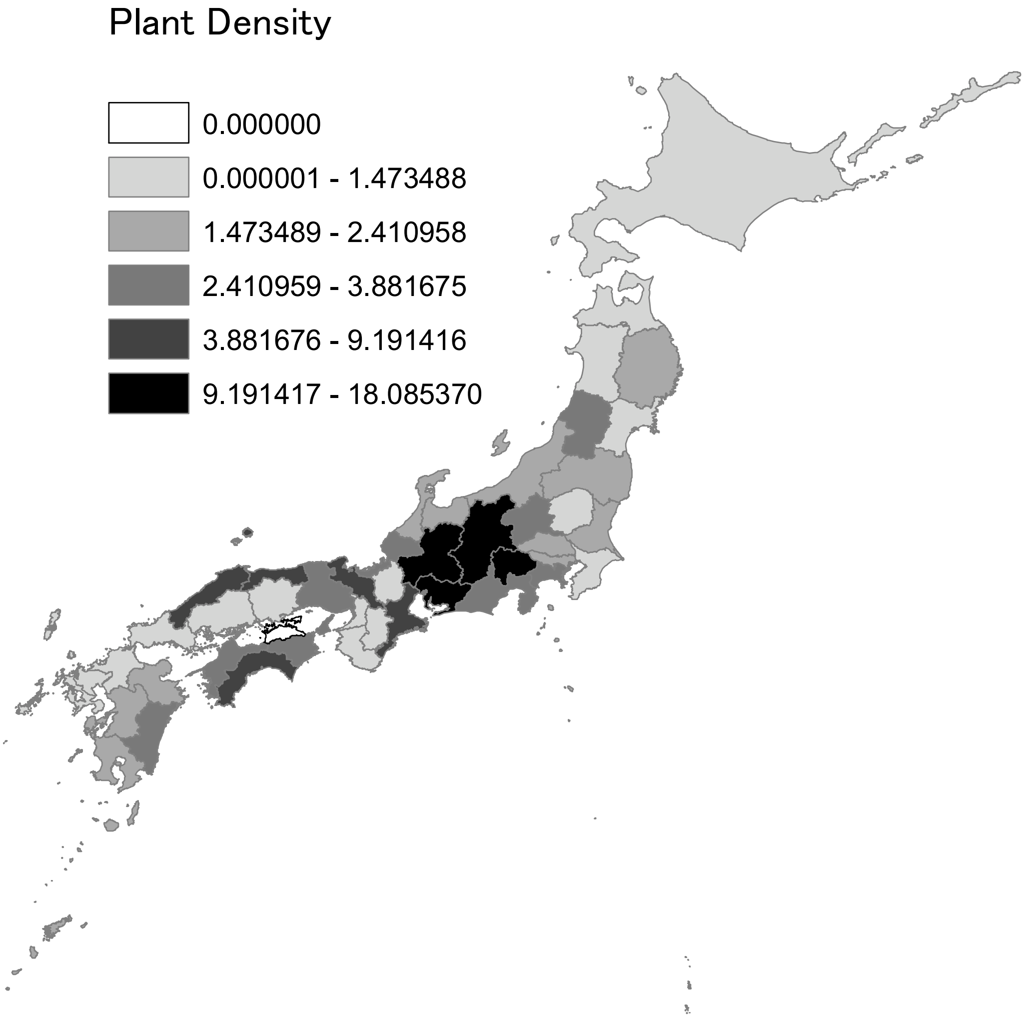
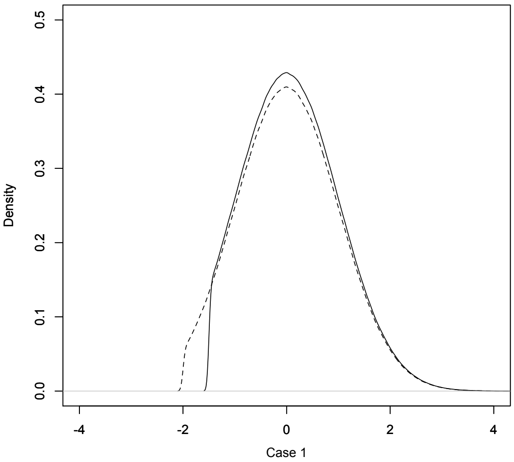
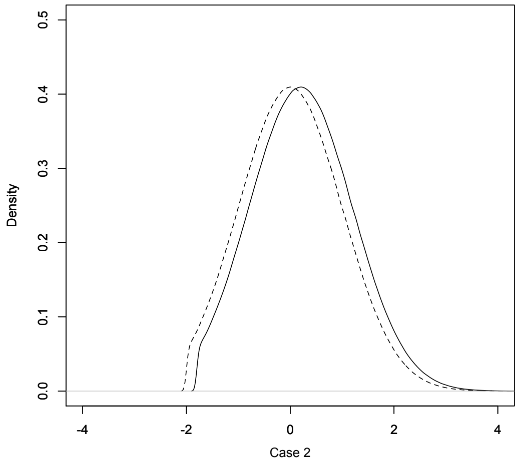


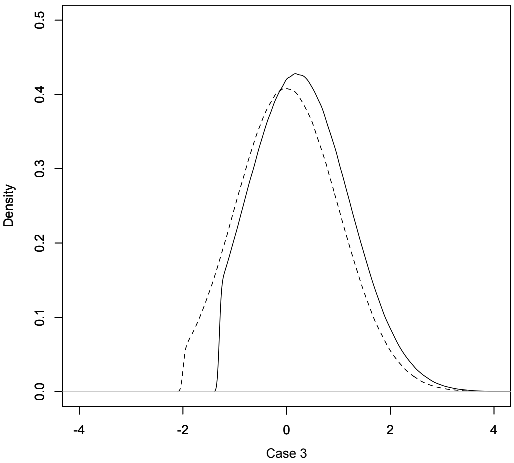

Figure 1: Map of Japan and the density of silk-reeling plants in 1909



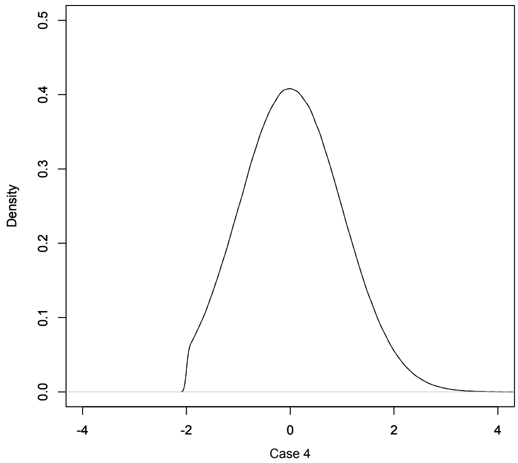

(a) Selection




(b) Agglomeration




(c) Both selection and agglomeration




(d) Neither effect

Figure 2: Four considerable cases of cluster effects

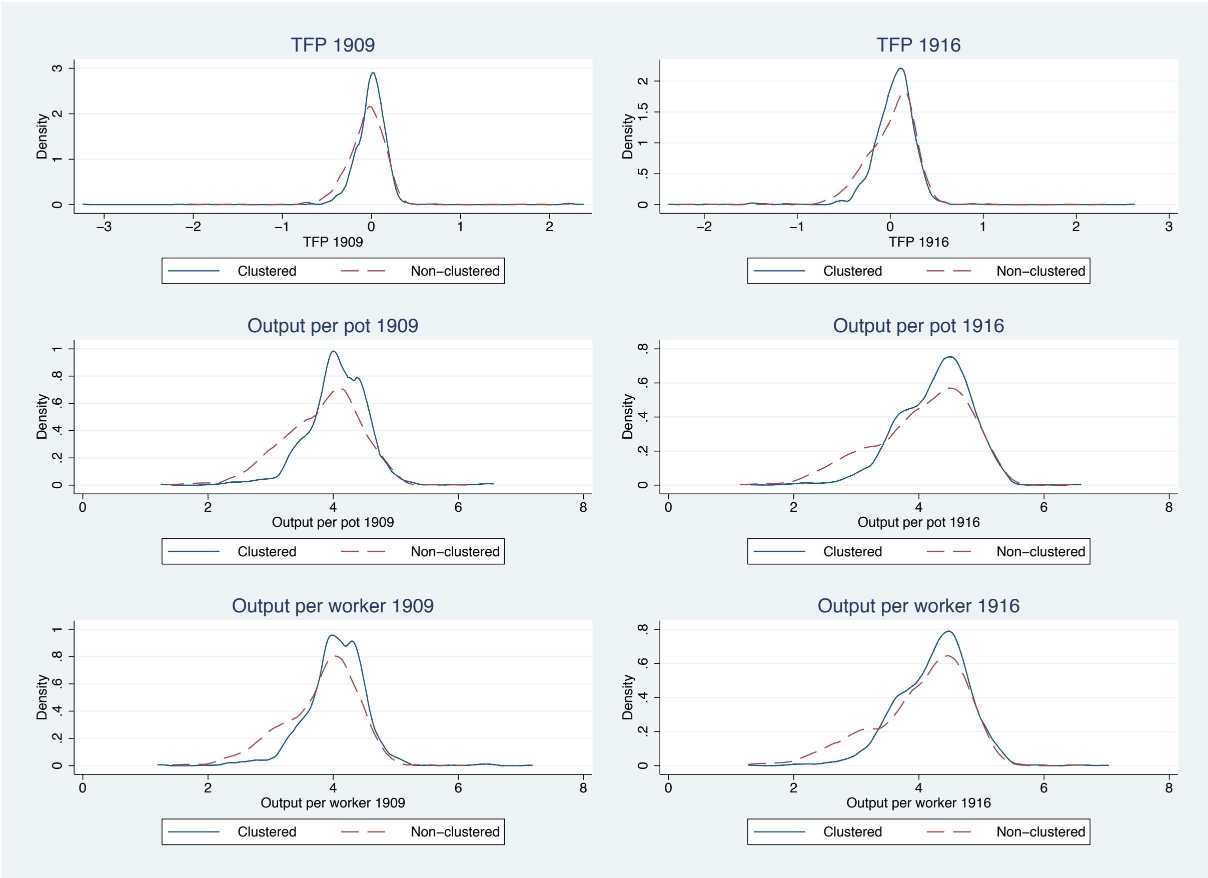



Figure 3: Kernel densities of plant-level productivity

# A Estimating TFP using an alternative method

As discussed in Section 4.2, we estimate TFP by the fixed effect estimation, assuming that the firm observable productivity is unchanged in the relevant period (seven years). To relax this assumption, we estimate an alternative measure of TFP using the method in Levinsohn and Petrin (2003) and check the robustness of our results to the alternative measure of TFP.

Again, our specification of the firm-level production function is as follows:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it}, \tag{23}$$

where $y_{it}$ is the log of output, $k_{it}$ is the log of capital input (number of pots), $l_{it}$ is the log of labor inputs (number of female workers), $m_{it}$ is the log of intermediate input (quantity of cocoons used), $\omega_{it}$ and $\epsilon_{it}$ are the productivity terms that are unobservable to the econometrician. While the $\epsilon_{it}$ are not also observed by firms before they make their input decisions, $\omega_{it}$ is the productivity measure observed by the plant. In the main analysis, we assume $\omega_{it} = \omega_i$ and estimate this equation using fixed effect estimation. Alternatively, this section estimates this production function by the method in Levinsohn and Petrin (2003), which does not assume $\omega_{it} = \omega_i$ but that $m_{it} = m_{it}(k_{it}, \omega_{it})$, and this intermediate input function is monotonically increasing in $\omega_{it}$ and invertible by $\omega_{it}$. Using the TFP estimated by the method in Levinsohn and Petrin (2003), this section replicated the analysis as the main analysis and check the robustness of our main conclusion.

The estimated TFP in this approach is highly correlated with the estimated TFP in the main analysis. The correlation between fixed effect OLS TFP and Levinsohn-Petrin TFP is 0.80. Descriptive statistics of Levinsohn-Petrin TFP are shown in Table 9.

= Table 9 =

Overall, the results are quite similar to the main results. First, both in 1909 and 1916, the interquantile range in the cluster is clearly shorter than that in the non-cluster. This suggests truncation of the distribution in cluster. Second, in lower tails (the 10th and 25th percentiles), the percentiles in clusters are much higher than in non-clusters. On the other hand, in higher tails, the percentiles in both the cluster and non-cluster are quite similar (the 90th percentile in the cluster and non-cluster are 3.390 and 3.322). Those two results are quite similar to the main analysis, which used the productivity measure estimated by fixed effect analysis.

Furthermore, to check the robustness of our main conclusion, we show the results of two key analyses. The first is the truncation effects analysis. We conduct the same analysis in Table 4 in Section 4.3. This is shown in Table 10.

= Table 10 =

Eq.(20) and eq.(21) are the estimation equations. Even by using alternative an measure of productivity, we observe a clear truncation of the productivity distribution. Column (1) represents the results of estimating eq. (20). The coefficient of the ln(Density) is negatively significant. This suggests truncation of the productivity distribution. Furthermore, coefficients of $\ln(\mathrm{D}_{pt})$ are positive and significant for the lower tail, that ism the 10th and 25th percentile points (columns 2 and 3). On the other hand, the coefficients of plant density are not statistically different from zero for the higher tail (the 90th percentile) in column (4). In the mid-tail (the 75th percentile), the coefficients of $\ln(\mathrm{D}_{pt})$ are positive and significant, but its absolute value is much smaller than the coefficient in the lower tails, and its p-value is relatively small (8.5 %). Thus, the agglomeration force that shifts distribution rightwards would be small. The results (i) left-truncation of the productivity distribution and (ii) little evidence of the rightward shift of the distribution are the same as the main results in Section 4.3.

The second key analysis is the growth analysis. We conduct the same analysis in Table 7 in Section 5.1. This is shown in Table 11.

= Table 11 =

Eq.(22) is the estimation equation. As in the main results in Section 5.1, in every specification, the coefficient of ln(Density) is not significant at the conventional levels, and the coefficient of the level of productivity in 1909 is negatively significant (columns 1 and 2). This holds even for restricting samples to start-up plants (columns 3 and 4). This suggests that there is no evidence that plants located in the cluster were learning faster and catching up with leading plants more than plants in non-clusters. Furthermore, these results hold even if we restrict samples to start-up plants (plants younger than five years). These results are the same as the main results in Section 5.1.

Overall, our key results hold even with the application of an alternative measure of productivity. That is, (i) productivity distribution in clusters is left truncated and there is no rightward shift of distribution in the cluster and (ii) after start-up, the plants located in clusters were not learning faster and catching up with leading plants more than the plants in non-clusters. Thus, our main conclusion was robust in the measure of productivity.

Table 9: Distribution of alternative measure of total factor productivity

| Panel A: 1909 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP (Levinsohn & Petrin) | All | 2.884 | 0.484 | 2.349 | 2.653 | 2.936 | 3.137 | 3.368 |
| | Cluster | 2.978 | 0.394 | 2.558 | 2.788 | 2.988 | 3.182 | 3.390 |
| | Non-cluster | 2.790 | 0.581 | 2.181 | 2.515 | 2.843 | 3.096 | 3.322 |
| | | (10.208) | | | | | | |
| Panel B: 1916 | | | | | | | | |
| Measure of Productivity | Cluster vs. Non-cluster | Mean | Interquartile range | 10th Percentile | 25th Percentile | 50th Percentile | 75th Percentile | 90th Percentile |
| TFP (Levinsohn & Petrin) | All | 3.019 | 0.657 | 2.319 | 2.722 | 3.110 | 3.379 | 3.582 |
| | Cluster | 3.082 | 0.608 | 2.500 | 2.789 | 3.140 | 3.397 | 3.583 |
| | Non-cluster | 2.938 | 0.762 | 2.112 | 2.588 | 3.053 | 3.350 | 3.577 |
| | | (6.548) | | | | | | |

Note: This table shows descriptive statistics of the plant-level productivities. TFP is estimated by the method in Levinsohn and Petrin (2003). The $t$-values of the $t$-test are in the parentheses. The null hypotheses of the $t$-test is that the average productivities are not different for clusters and non-clusters.

Table 10: Truncation effects in alternative measure of productivity

| Dependent variable | (1) Interquartile range | (2) 10th percentile | (3) 25th percentile | (4) 75th percentile | (5) 90th percentile |
|---|---|---|---|---|---|
| ln(Density) | -0.0766** | 0.147** | 0.146** | 0.0693* | 0.0622 |
| | (0.0313) | (0.0589) | (0.0540) | (0.0371) | (0.0392) |
| Constant | 0.603** | 2.265** | 2.486** | 3.088** | 3.277** |
| | (0.0573) | (0.103) | (0.0938) | (0.0587) | (0.0604) |
| Year fixed effect | yes | yes | yes | yes | yes |
| Observations | 45 | 45 | 45 | 45 | 45 |
| Adjusted $R^2$ | 0.121 | 0.104 | 0.128 | 0.222 | 0.248 |

Note: Column (1) shows estimated coefficients on the regression of the interquartile range of the plant-level productivity distribution in prefectures on plant density. The dependent variables are the interquartile range of the productivity distribution at the prefectural level. Columns (2) to (5) show estimated coefficients on the regression of percentile points of the plant-level productivity distribution in prefectures on plant density. The dependent variables are percentile points of the productivity distribution at the prefectural level. TFP is estimated by the method in Levinsohn and Petrin (2003).
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

Table 11: Estimation of growth effects in alternative measure of productivity

| Dependent variable | (1) TFP growth | (2) TFP growth | (3) TFP growth | (4) TFP growth |
|---|---|---|---|---|
| ln(density) | -0.0148 | -0.00968 | -0.0589 | -0.0911 |
| | (0.0284) | (0.0258) | (0.0919) | (0.0832) |
| ln(pot) | 0.968** | 0.0658 | 2.393** | 1.418** |
| | (0.393) | (0.418) | (0.438) | (0.586) |
| ln(worker) | -0.340 | -0.205 | -0.816 | -0.737 |
| | (0.418) | (0.392) | (0.638) | (0.580) |
| ln(cocoon) | -0.552** | 0.0711 | -1.472** | -0.707 |
| | (0.183) | (0.134) | (0.554) | (0.497) |
| ln(age) | -0.0783* | -0.0719* | -0.182 | -0.168 |
| | (0.0415) | (0.0395) | (0.222) | (0.197) |
| Steam power dummy | 0.0531 | 0.0978 | 0.398 | 0.374 |
| | (0.0791) | (0.0755) | (0.262) | (0.257) |
| Water power dummy | -0.0316 | -0.0148 | 0.701* | 0.730** |
| | (0.0811) | (0.0785) | (0.355) | (0.347) |
| ln(TFP) in 1909 | | -1.100** | | -1.114** |
| | | (0.203) | | (0.483) |
| Constant | 1.370** | 3.723** | 2.870** | 4.818** |
| | (0.420) | (0.639) | (1.225) | (1.706) |
| Observations | 909 | 909 | 198 | 198 |
| Adjusted $R^2$ | 0.157 | 0.265 | 0.534 | 0.574 |

Note: This table shows the estimated coefficients of the regression of productivity growth on plant density. The dependent variables are the growth rates of the plant-level productivity from 1909 to 1916. TFP is estimated by the method in Levinsohn and Petrin (2003).
Robust standard errors in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.

# B Mathematical Appendix (not for publication)

## B.1 Derivation of $\frac{d\hat{h}}{ds}$ (eq.(9))

The implicit function theorem implies

$$\frac{d\hat{h}}{ds} = \frac{-(\partial V^e/\partial s)}{\partial V^e/\partial \hat{h}}$$

$-(\partial V^e/\partial s) = 1$ is immediate from (8). The denominator is

$$\begin{aligned}
\frac{\partial V^e}{\partial \hat{h}} &= \left[\frac{\left(\hat{h} - \hat{h} + \sqrt{wf}\right)^2}{w} - f\right] g(\hat{h}) + 2\int_0^{\hat{h}} \left[\frac{\hat{h} - h + \sqrt{wf}}{w}\right] g(h)dh \\
&= 2\int_0^{\hat{h}} \left[\frac{\hat{h} - h + \sqrt{wf}}{w}\right] g(h)dh > 0
\end{aligned}$$

Therefore,

$$\frac{d\hat{h}}{ds} = \frac{-(\partial V^e/\partial s)}{\partial V^e/\partial \hat{h}} = \frac{w}{2\int_0^{\hat{h}} \left(\hat{h} - h + \sqrt{wf}\right) g(h)dh} > 0.$$

## B.2 Derivation of $N_e$ (eq.(13))

Total producution in a region is

$$Q = N_e \int_0^{\hat{h}} q_i^* g(h)dh.$$

Inserting $q_i^*$ from (4) yields:

$$\begin{aligned}
Q &= N_e \int_0^{\hat{h}} q_i^* g(h)dh \\
&= N_e \int_0^{\hat{h}} \frac{p - h - wE(Q)}{w} g(h)dh \\
&= N_e \int_0^{\hat{h}} \frac{p - h}{w} g(h)dh - N_e E(Q) G(\hat{h})
\end{aligned}$$

Since $Q = E(Q)$ in the equilibrium, we can write $Q$ as a function of $N_e$ and $\hat{h}$:

$$Q(N_e, \hat{h}) = \frac{N_e}{1 + N_e G(\hat{h})} \int_0^{\hat{h}} \frac{p - h}{w} g(h)dh.$$

Also, by rearranging (6) ($\hat{h} = p - wQ - \sqrt{wf}$), we get another expression of

$$Q = \frac{p - \hat{h} - \sqrt{wf}}{w}.$$

So, we have

$$
\begin{aligned}
&\frac{N_e}{1+N_e G(\hat{h})}\int_0^{\hat{h}} \frac{p-h}{w} g(h)dh = \frac{p-\hat{h}-\sqrt{wf}}{w} \\
\Rightarrow \;& N_e \int_0^{\hat{h}} (p-h)g(h)dh = [1+N_e G(\hat{h})](p-\hat{h}-\sqrt{wf}) \\
\Rightarrow \;& N_e \left[\int_0^{\hat{h}} (p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})G(\hat{h})\right] = p-\hat{h}-\sqrt{wf} \\
\Rightarrow \;& N_e \left[\int_0^{\hat{h}} (p-h)g(h)dh - \int_0^{\hat{h}} (p-\hat{h}-\sqrt{wf})g(h)dh\right] = p-\hat{h}-\sqrt{wf} \\
\Rightarrow \;& N_e \left[\int_0^{\hat{h}} (\hat{h}-h+\sqrt{wf})g(h)dh\right] = p-\hat{h}-\sqrt{wf}
\end{aligned}
$$

and hence

$$
N_e = \frac{p-\hat{h}-\sqrt{wf}}{\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh}.
$$

### B.3 Derivation of $\frac{\partial N_e}{\partial \hat{h}}$

$$
\begin{aligned}
\frac{\partial N_e}{\partial \hat{h}} &= \frac{-\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh - (p-\hat{h}-\sqrt{wf})\left[\sqrt{wf}g(\hat{h}) + \int_0^{\hat{h}} g(h)dh\right]}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} \\
&= \frac{-\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh - \int_0^{\hat{h}}(p-\hat{h}-\sqrt{wf})g(h)dh - (p-\hat{h}-\sqrt{wf})\sqrt{wf}g(\hat{h})}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} \\
&= \frac{-\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})\sqrt{wf}g(\hat{h})}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} < 0
\end{aligned}
$$

### B.4 Derivation of $dN_p/ds$

Since $N_p = N_e G(\hat{h})$,

$$
\frac{dN_p}{ds} = \frac{dN_e}{d\hat{h}}\frac{d\hat{h}}{ds}G(\hat{h}) + N_e g(\hat{h})\frac{d\hat{h}}{ds} = \left[\frac{dN_e}{d\hat{h}}G(\hat{h}) + N_e g(\hat{h})\right]\frac{d\hat{h}}{ds}
$$

By using (13) and (14), and letting

$$
\xi \equiv \int_0^{\hat{h}} (\hat{h}-h+\sqrt{wf})g(h)dh,
$$

the term in the brackets is

$$
\begin{aligned}
\frac{dN_e}{d\hat{h}}G(\hat{h}) + N_e g(\hat{h}) &= \frac{-G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})\sqrt{wf}g(\hat{h})G(\hat{h})}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} + \frac{(p-\hat{h}-\sqrt{wf})g(\hat{h})}{\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh} \\
&= \frac{-G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})\sqrt{wf}g(\hat{h})G(\hat{h}) + (p-\hat{h}-\sqrt{wf})g(\hat{h})\xi}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} \\
&= \frac{-G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})g(\hat{h})\left[\sqrt{wf}G(\hat{h}) - \xi\right]}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2} \\
&= \frac{-G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})g(\hat{h})\left[\sqrt{wf}G(\hat{h}) - \int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2}.
\end{aligned}
$$

Let $J$ denote the numerator.

$$
\begin{aligned}
J &= -G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh - (p-\hat{h}-\sqrt{wf})g(\hat{h})\left[\sqrt{wf}G(\hat{h}) - \int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right] \\
&= -G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}\left(\hat{h}-h\right)g(h)dh \\
&= -G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\hat{h}G(\hat{h}) - (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}hg(h)dh \\
&= -G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\hat{h}G(\hat{h}) - (p-\hat{h}-\sqrt{wf})g(\hat{h})\left[\hat{h}G(\hat{h}) - \int_0^{\hat{h}}G(h)dh\right] \\
&= -G(\hat{h})\int_0^{\hat{h}}(p-h)g(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}G(c)dh \\
&= -pG(\hat{h})^2 + G(\hat{h})\int_0^{\hat{h}}hg(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}G(h)dh \\
&= -pG(\hat{h})^2 + G(\hat{h})\left(\hat{h}G(\hat{h}) - \int_0^{\hat{h}}G(h)dh\right) + (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}G(h)dh \\
&= -pG(\hat{h})^2 + \hat{h}G(\hat{h})^2 - G(\hat{h})\int_0^{\hat{h}}G(h)dh + (p-\hat{h}-\sqrt{wf})g(\hat{h})\int_0^{\hat{h}}G(h)dh \\
&= -(p-\hat{h})G(\hat{h})^2 - \left[G(\hat{h}) + \sqrt{wf}g(\hat{h})\right]\int_0^{\hat{h}}G(h)dh + (p-\hat{h})g(\hat{h})\int_0^{\hat{h}}G(h)dh \\
&= -(p-\hat{h})\left[G(\hat{h})^2 - g(\hat{h})\int_0^{\hat{h}}G(h)dh\right] - \left[G(\hat{h}) + \sqrt{wf}g(\hat{h})\right]\int_0^{\hat{h}}G(h)dh.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\frac{dN_p}{ds} &= \left[\frac{dN_e}{d\hat{h}}G(\hat{h}) + N_e g(\hat{h})\right]\frac{d\hat{h}}{ds} \\
&= \frac{-(p-\hat{h})\left[G(\hat{h})^2 - g(\hat{h})\int_0^{\hat{h}}G(h)dh\right] - \left[G(\hat{h}) + \sqrt{wf}g(\hat{h})\right]\int_0^{\hat{h}}G(h)dh}{\left[\int_0^{\hat{h}}(\hat{h}-h+\sqrt{wf})g(h)dh\right]^2}\frac{d\hat{h}}{ds} < 0. \qquad (24)
\end{aligned}
$$

This is negative because $\left[G(\hat{h})^2 - g(\hat{h})\int_0^{\hat{h}} G(h)dh\right]$ is positive, since $G(\hat{h}) \geq G(h)$ for all $h \in [0, \hat{h}]$ and $G(\hat{h}) > g(\hat{h})$,