RIETI Discussion Paper Series 10-E-056

Spatial Relocation with Heterogeneous Firms and Heterogeneous Sectors

OKUBO Toshihiro
Kobe University

Rikard FORSLID
University of Stockholm

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

RIETI Discussion Paper Series 10-E-056

October 2010

Spatial Relocation with Heterogeneous Firms and Heterogeneous Sectors*

Toshihiro Okubo

Kobe University

Rikard Forslid

University of Stockholm

Abstract

The present paper focuses on sorting as a mechanism behind the well-established fact that there is a central region productivity premium. Using a model of heterogeneous firms that can move between regions, Baldwin and Okubo (2006) show how more productive firms sort themselves to the large core region. We extend this model by introducing different capital intensities among firms and sectors. In accordance with empirical evidence, more productive firms are assumed to be more capital intensive. As a result, our model can produce sorting to the large regions from both ends of the productivity distribution. Firms with high capital intensity and high productivity, as well as firms with very low productivity and low capital intensity, tend to relocate to the core. We use region and sector productivity distributions from Japanese micro data to test the predictions of the model. Several sectors show patterns consistent with two-sided sorting, and roughly an equal number of sectors seem to primarily be driven by sorting and selection. We also find supportive evidence for our model prediction that two-sided sorting occurs in sectors with high capital intensity.

Keywords: Agglomeration; Firm Heterogeneity; Productivity, Spatial Sorting

JEL classification: F12, F15, F21, R12

* This research was conducted as part of the project on "Study on Offshore Outsourcing by Japanese Firms" at the Research Institute of Economy, Trade and Industry (RIETI).

1 Introduction

Heterogeneity on the supply side is currently in vogue in many fields of economics such as macroeconomics, international trade and economic geography, where micro datasets make it possible to study the behaviour of individual firms. Models of heterogeneous firms are used in the international trade literature to explain observed differences between exporters and non-exporters in terms of e.g. size and productivity (see Melitz 2003, Helpman et al. 2004 and Melitz and Ottaviano 2008). However, while the study of firms during the current globalization period is highly important, it has lead to a relative neglect of a traditional focus of trade theory; namely that of heterogeneous sectors (see e.g. Neary 2009).[1]

The present paper analyses a setting with heterogeneous sectors and heterogeneous firms. We focus on the impact of sector and firm heterogeneity on firm location and on the effects on the firm productivity distribution of different locations. It is empirically well established in the urban economics and economic geography literature that firms in core areas, such as urban areas or densely populated manufacturing areas, tend to be more productive than those in peripheral regions (see Rosenthal and Strange 2004 and Melo et al. 2009 for a survey). Common explanations for this empirical finding are positive agglomeration externalities related to technological spillovers, labour market pooling or better access to suppliers and customers. Another source of higher productivity in core locations is stronger selection among firms in the core, as pointed out in the heterogenous firms literature (see Melitz 2003, Helpman et al. 2004, and Melitz and Ottaviano 2008). Stronger competition in larger markets will induce the least efficient firms to close down, thereby increasing average productivity.[2] A third mechanism, which is the focus of the present paper, is spatial sorting of heterogeneous firms. Baldwin and Okubo (2006) show how high productivity firms would tend to sort themselves to the larger regions.[3] Their theoretical framework combines the 'footloose capital' trade and location model by Martin and Rogers (1995) and the heterogeneous firms model by Melitz (2003). Spatial sorting happens because more productive firms have higher sales and therefore have more to gain from lower transportation costs in the large core market. They are also better equipped for coping with the higher competition in the core.

The present paper introduces different capital (or R&D) intensities among firms *and* sectors in the spatial sorting model by Baldwin and Okubo (2006). In line with empirical evidence, we assume that large and highly productive firms use more capital (R&D) and relatively less labour. We also allow the tendency to higher capital intensity among more productive firms to

[1] Some important exceptions exist, such as Bernard, Redding, and Schott (2007) who analyse sectors of different capital labour ratios in a heterogeneous firms model.

[2] The differential selection effect in small and large markets is present in the original Melitz (2003) model as well as in Melitz and Ottaviano (2005). However, many other models simplify away the market size effect e.g. by assuming that wages are equalised by free trade in some constant returns sector.

[3] Okubo and Tomiura (2010b) tested this hypothesis, and found that low productivity firms are more likely to relocate from the core to the periphery in respons to a regional subsidy.

be sector specific. The capital intensity (or R&D intensity) differs substantially among firms and sectors in the data. Large and highly productive firms use more capital, have a high R&D intensity and use relatively less labour.[4] At the sectoral level, chemical, pharmaceutical and machinery industries are among the most highly capital and R&D intensive sectors.[5] The assumption of higher capital intensity among more productive firms implies that our model can generate sorting from both ends of the productivity distribution (two-sided sorting). Firms with the highest return to capital have the strongest incentive to move from the periphery to the core region. In our setting, these are the most productive firms as well as the least productive firms that have a very high labour to capital ratio. An implication of this is that the model can generate core regions that have the most productive and mechanised firms as well as e.g. firms producing high priced hand-made items. We allow the tendency to two-sided sorting to be sector-specific in our model, which is consistent with empirical evidence showing very different location patterns among sectors (see e.g. Combe and Overman 2002).

While agglomeration externalities, selection and sorting all produce higher average productivity among firms in the core, they have very different implications for the second- and third-order moments of the productivity distribution of firms. Agglomeration externalities imply a upward shift of the entire distribution, which implies higher average productivity but unchanged variance and skewness. Selection implies a truncation at the low end of the distribution in the core, as the least productive firms are forced out of business. This implies that the productivity distribution in the core has a lower variance (see Gatto, Ottaviano and Paganini, 2008). It also implies negative skewness in the core. Sorting, in contrast, would lead to a higher variance (spread) in the core, as firms from the end(s) of the firm productivity distribution in the periphery move to the core. Also one-sided sorting implies positive skewness in the periphery, but this effect is dempened when there is two-sided sorting.

A few papers have used firm-level data to test for selection effects on firm productivity or cost distributions. International trade implies stronger competition and is therefore one factor that would lead to stronger firm selection. Syverson (2004) does not find any relationship between spreads of the productivity distribution of firms and tradeability using a cross-section of U.S. firms. In contrast, using a panel of Italian firms, Gatto, Ottaviano and Paganini (2008) find that intraindustry cost spreads are smaller in export oriented industries. Combes et al. (2009) use a quantile regression on firm establishment data to establish the relative importance of agglomeration versus selection for the size of the productivity premia related to French cities. They find that spatial productivity differences are mainly explained by agglomeration but that selection is important for some relatively disaggregated sectors.

[4] One explanation for this could be that due to financial constraints, small firms have difficulties in financing capital investment (see Cabral and Mata 2002).

Hall (1992) points out the realtionship between R&D financing and firm size. Boothby et al. (2008) and Cohen and Klepper (1996) show that R&D expenditures are proportional to firm size.

[5] For example, 60 to 70 percent of the total R&D expenditures in manufacturing are spent by the machinery sectors only, according to Japanese sectoral data (JIP data in RIETI).

This paper instead attempts to identify sectors where sorting is important. For this porpose, we use firm plant level data from Japan's *Census of Manufacturers* covering virtually all plants with more than five employees in 1990 classified at the three-digit sector level. We estimate region- and sector-specific kernel density functions for productivity, and we find that a large number of sectors display productivity distributions consistent with one- or two-sided sorting, and likewise that many sectors are consistent with selection and agglomeration. We also find supportive evidence for our model prediction that two-sided sorting ocurrs in sectors with a high capital intensity.

In a purely empirical paper, Okubo and Tomiura (2010a) use the same dataset to estimate the aggregate productivity distribution on a regional level. They find a productivty premium in the core, but also that the core hosts some low-productivty firms. This finding is consistent with the present paper.

Our paper is related to that of Okubo, Picard and Thisse (2010) which uses the linear-demand monopolistic competition set-up of Ottaviano, Tabuchi and Thisse (2002) to analyse the location choice of two types of firms: low productive and high productive. Because of lower mark-ups due to tougher competition in the large market, only the most productive firms will initially survive in that market. Competition spreads to both regions as trade costs come down enough, which also leads the low productivity firms to prefer the large market. This outcome has similarities to our two-sided sorting equilibrium. However, our results are driven by a completely different mechanism where firm and sector differences in capital intensity play a crucial role. Our empirial analysis using micro data also supports the notion that two-sided sorting is related to the capital intensity of sectors.

The next section presents the model. Section 3 contains empirical analysis. Finally, section 4 concludes the paper.

2 The Model

This chapter uses the Baldwin and Okubo (2006) heterogenous firm version of the 'footloose capital' new economic geography model by Martin and Rogers (1995). The model is enriched by allowing for different capital intensities among firms. It is assumed that higher productivity is associated with a higher capital stock, as documented by numerous empirical studies on micro data (see Bernard et al. 2007).

2.1 Basics

There are two regions with assymmetric population (or market size). Core is the large region and Periphery (denoted by *) is the small region. There two types of factors of production, capital and labour. Capital, which is sector specific, can move between regions but capital owners do not. Workers can move freely between sectors but are immobile between regions. The larger region, Core, is endowed with the share $s(> 0.5)$, and the smaller region, Periphery, with $1 - s$

of the world endowment of labour and capital, that is, countries are of different size, but they have identical capital labour ratios. A homogeneous good is produced with a constant-returns technology only using labour. Differentiated manufactures are produced with increasing-returns technologies using both capital and labour. There are m sectors of differentiated goods. The mass of firms in each sector is normalised to one, $N_m \equiv 1$, which means that the home country has s firms in each sector at outset. Firm productivities in each sector are distributed according to a cumulative density function, $F_m(a)$.

The firms' productivity level is also associated with firm-specific capital requirement. It is assumed that more productive firms have a higher capital requirement. However, this relationship is sector specific. All individuals have the utility function

$$U = C_M^{\mu} C_A^{1-\mu}, \quad where \quad C_M = \prod_m C_m^{\theta_m} \tag{1}$$

where $\mu \in (0,1)$ and $\theta_m > 0$ are constants, and country subscripts are suppressed for ease of notation $\sum \theta_m = 1$, C_A is consumption of the homogenous good and differentiated goods from each manufacturing sector enter the utility function through a sector-specific index C_m, defined by

$$C_m = \left[\int_{k \in \Psi} c_{km}^{(\sigma-1)/\sigma} dk \right]^{\sigma/(\sigma-1)}, \tag{2}$$

Ψ being the set of varieties consumed, c_{km} the amount of variety k from sector m consumed, and $\sigma > 1$ the elasticity of substitution.

Each consumer spends a share μ of his income on manufactures, and constant fractions θ_m of this are spent on varieties from each sector. Thus, it is possible to separately analyse the equilibrium for each sector, and therefore we will henceforth when possible suppress the sector indices. Total demand for a domestically produced variety i in a sector m is

$$x_i = \frac{p_i^{-\sigma}}{\int_{k \in \Psi} p_k^{1-\sigma} dk} \cdot \theta_m \mu Y, \tag{3}$$

where p_k is the price of variety k, and Y income in the region.

The unit factor requirement of the homogeneous good is one unit of labour. This good is freely traded, and since it is also chosen as the numeraire, we have

$$p_A = w = 1, \tag{4}$$

w being the wage of workers in both regions.

Ownership of capital is assumed to be fully interregionally diversified; that is, if one region owns X-percent of the world capital stock, it will own X-percent of the capital in each region. The income of each region is therefore constant and independent of the location of capital. World expenditure equals world factor income $E^W = wL^W + \sum \theta_m \mu E^W / \sigma$. Without loss of generality we choose units so that $L^W \equiv 1$, which gives $E^W = \frac{1}{1-\mu/\sigma}$.Income in Core is equal to its share of world expenditures given by

$$Y = s\mu E^W = s\frac{\sigma}{\sigma - \mu}. \tag{5}$$

Y is thus constant irrespective of the location of capital; i.e. also out of long-run equilibrium.

In the production of differentiated goods, the fixed cost consists of capital, whereas the variable cost consists of labour. Firms are differentiated, and their firm-specific marginal production costs a_i are distributed according to the cumulative distribution function $F(a)$. Here, it is also assumed that firms with a lower a has a higher capital cost.[6] The capital requirement for a firm with the labour input coefficient a in sector m is given by $h_m(a)$, which is a concave function in a. Importantly sectoral heterogeneity in our model is simply expressed by differences in $h_m(a)$, and we write out the sector subscript to stress this. The underlying motivation for having different and sector specific h functions is the above mentioned fact that capital intensity and capital requirement are substantially heterogeneous across sectors as well as across firms.

Distance is represented by trading costs. Shipping the manufactured good involves a frictional trade cost of the "iceberg" form: for one unit of good from region j to arrive in region l, $\tau_{jl} > 1$ units must be shipped. Trade costs are also assumed to be equal in both directions so that $\tau_{jl} = \tau_{lj}$.

Profit maximisation by manufacturing firms leads a constant mark-up over marginal cost

$$p_i = \frac{\sigma}{\sigma - 1} a_i, \tag{6}$$

2.2 Short-run equilibrium

Similar to Baldwin and Okubo (2006), a_i is ramdomely allocated among firms. However, different from that model, our model involves two factors (capital and labour), and different $a_i' s$ create both heterogeneous capital requirements and per-unit labour requirements. In the short run, the allocation of K^W is taken to be fixed. In order to solve the model analytically, we follow Helpman, Melitz and Yeaple (2004) and assume the cumulative density function of a to be Pareto[7]:

$$F(a) = \left(\frac{a}{a_0}\right)^{\rho}, \tag{7}$$

where $\rho > 1$ is a shape parameter and a_0 is a scaling parameter. Without loss of generality we assume that $a_0 = 1$. Figure 1 illustrates the distribution of firms in the two economies before capital can move.

We also assume the following simple relationship between a_i and the fixed capital requirement:

[6] This is a standard finding among micro data studies. See e.g. Bernard et al. (2007).

[7] This assumption is consistent with the empirical findings by e.g. Axtell (2001) or Luttmer (2007).

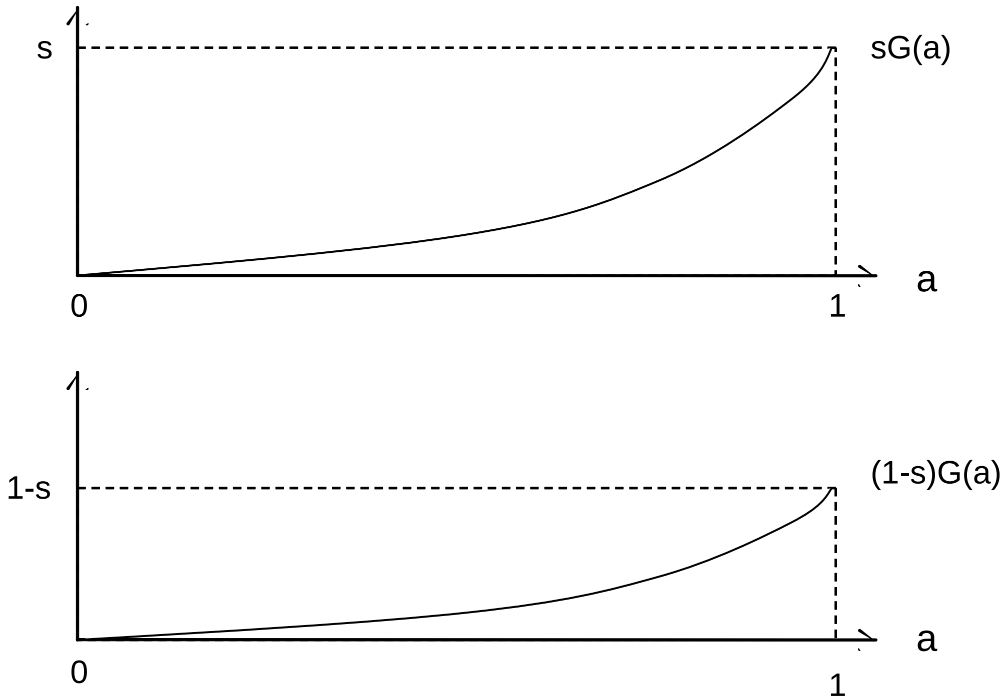

Figure 1: The initial distribution of firm in the two economies

$$h_m(a_i) = 2 - a_i^{\gamma_m}, \tag{8}$$

where importantly $\gamma_m \geq 0$ is a sector-specific parameter. For $\gamma_m = 0$ we obtain the standard footloose capital model, while sectors with a positive γ have increasing capital requirements for high productivity firms. We interpret sectors with a high γ as sectors with important scale economies related to fixed investments in e.g. R&D.[8]

A Core firm's return to capital is the operating profit divided by the firm's capital stock,

$$\pi(a_i) = \frac{a_i^{1-\sigma}}{h(a_i)(\sigma-\mu)}\theta\mu\left(\frac{s}{\overline{\Delta}} + \frac{\phi(1-s)}{\overline{\Delta}^*}\right), \tag{9}$$

where sector indices are suppressed, the right-hand side follows from the demand functions in (3), and where

$$\overline{\Delta} \equiv s\int_0^1 a_i^{1-\sigma}dF(a) + (1-s)\phi\int_0^1 a_i^{1-\sigma}dF(a), \tag{10}$$

$$\overline{\Delta}^* \equiv \phi s\int_0^1 a_i^{1-\sigma}dF(a) + (1-s)\int_0^1 a_i^{1-\sigma}dF(a). \tag{11}$$

[8] The capital stock corresponding to $N_m = 1$ is given by $K_m = \int_0^1 h_m dF(a) = 2 - \frac{\rho}{\gamma_m+\rho}$.

We note that the number of firms in each sector $N_m = 1$, which implies that the mass of firms in the Core is s. The object $\phi_{jl} = \tau_{jl}^{1-\sigma}$, ranging between 0 and 1, stands for "freeness" of trade between countries j and l (0 is autarchy and 1 is zero trade costs). It is assumed that the labour stock is sufficiently large so that the agricultural sector, which pins down the wage, is active in all regions.

Consider now what would happen if firms were allowed to move between regions. If all firms have unit capital requirements ($\gamma_m = 0$) as in the standard FC-model, the firms' return to capital, $\frac{a_i^{1-\sigma}}{(\sigma-\mu)}(1-\phi)\theta_m\mu\left(\frac{s}{\Delta} - \frac{1-s}{\Delta^*}\right)$ is convex and falling in a_i. Firms with the highest labour productivity (lowest a_i) will have the strongest incentives to move to the large market. Under reasonable assumptions of moving costs, this would lead to sorting with the most productive firms in the larger market, as shown by Baldwin and Okubo (2006). In the present paper, on the contrary, more efficient firms need more capital and the effect of a, on the return to capital, depends on the term $\frac{a_i^{1-\sigma}}{h_m(a_i)}$. Since $h_m(a_i)$ is concave and $a_i^{1-\sigma}$ is convex in a_i it will, under certain conditions, be the case that return to capital is highest for firms with a low a_i *and* firms with a high a_i. Intuitively, firms with the highest sales per unit of capital is either firms with a very high productivity or firms with a very high labour to capital ratio. These firms are then also the firms with the strongest incentives to move to the larger region in our model, once we allow capital to move.

More formally, a firm will move from the periphery to the core when

$$\pi(a_i) - \pi^*(a_i) - \chi = \frac{a^{1-\sigma}}{(\sigma-\mu)(2-a_i^{\gamma_m})}(1-\phi)\mu\theta_m\left(\frac{s}{\Delta} - \frac{1-s}{\Delta^*}\right) - \chi \geqslant 0, \tag{12}$$

where χ is a per-firm fixed moving cost. Once firms relocate between countries, moving costs are required to pay. The shape of this function is determined by the term $\frac{a^{1-\sigma}}{(2-a_i^{\gamma_m})}$, and it is easily shown by differentiation that it is U-shaped in a under the condition that $\sigma - 1 < \gamma_m$.

Figure 2 illustrates the U-shaped return to capital differential. Firms between a_L and a_H in the figure will tend to locate in the Periphery, since the gains from moving are higher than the moving cost χ, while firms at the ends of the productivity distribution to the right of a_U and to the left of a_L locate in the large region.

The U-shaped curve is shifted by e.g. changes in trade costs. At the point where the return to capital differential curve is tangent to the moving cost χ, and $a_L = a_U$, the model generates full agglomeration with all manufacturing firms located in the large region, which is the sustain point. The "a" with $a_L = a_U$ at the sustain point, denoted by a^S, is where the U-curve has its minimum and using (12), it gives the marginal cost at this point $a^S = \left(\frac{2(\sigma-1)}{\sigma-1+\gamma_m}\right)^{\frac{1}{\gamma_m}}$. Plugging this back into (12) gives the moving cost at which all industry moves to the core

$$\chi^S = \frac{(2(\sigma-1))^{\frac{1-\sigma}{\gamma_m}}}{2\gamma_m}(\sigma-1+\gamma_m)^{\frac{\gamma_m-1+\sigma}{\gamma_m}}\frac{\mu\theta_m B(1-\phi)}{\sigma-\mu}, \tag{13}$$

where $B \equiv \left(\frac{s}{\Delta} - \frac{(1-s)}{\Delta^*}\right)$. It is also relatively straightforward to see that

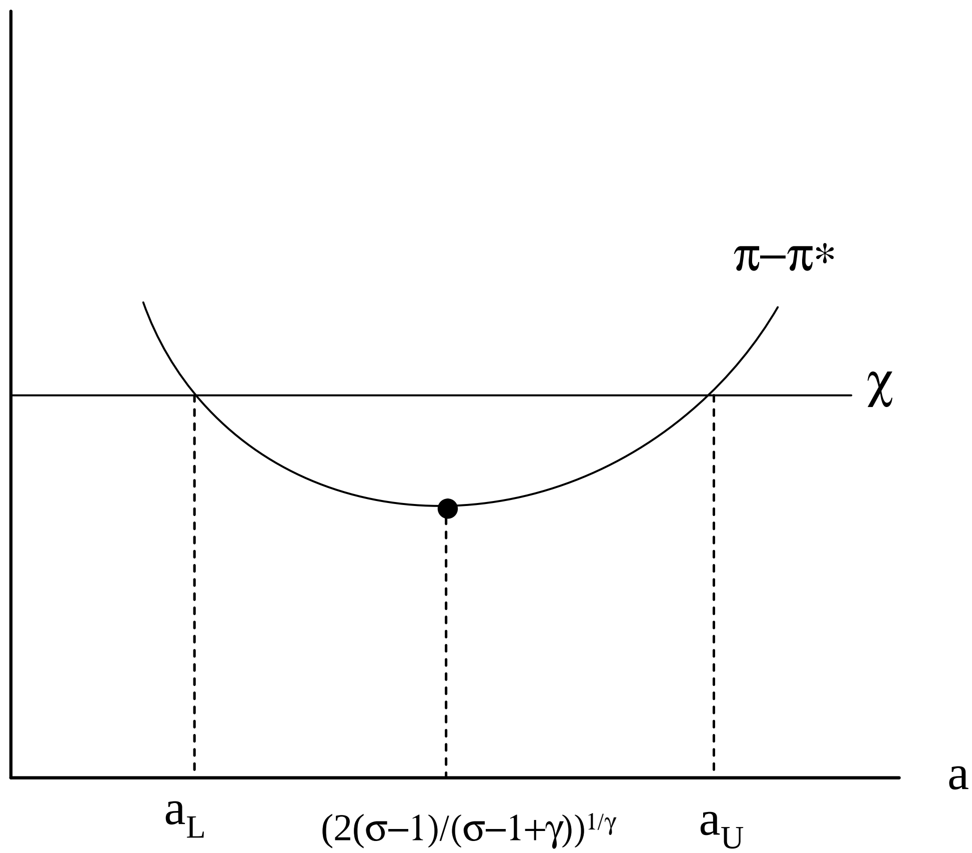

Figure 2: U-shaped return to capital differentials

$$\frac{\partial \chi^S}{\partial \gamma_m} < 0. \tag{14}$$

That is, sectors with a high γ_m will require lower moving costs before full agglomeration occurs. The expression (13) also implicitly defines trade freeness at the sustainpoint. Because the home-market effect is hump-shaped in trade costs, so will χ^S be.

The analysis so far has described the locational forces affecting firms in the short run. However, to analyse the long run when firms start to move, we need to explicitly model the dynamics, and we turn to this in the next section.

2.3 Long-run equilibrium

In the long run, capital is fully mobile between regions and responsive to the incentives provided by the relative returns that can be attained in the two countries.[9] We assume a moving cost that increases with the flow of migrating firms from the smaller region to the larger one, essentially a congestion cost, but we also introduce a fixed moving cost χ that is independent of the migration flow.[10] There is no discounting. The migration pressure will push up migration costs so that the first firm to migrate is the firm with the largest gains from this, namely the firm with the lowest a (see Baldwin and Okubo 2006). The value of migrating for a firm, v, in general depends on its own marginal cost and the mass of firms that have already migrated from each side of the distribution, a_{LR} a_{UR}. The value of migrating at a point in time for firms from the upper and lower part of the distribution is therefore

$$\begin{aligned} v_U(a_{UR}, a_{LR}) &= \pi(a_{UR}) - \pi^*(a_{UR}) - \chi \\ &= \frac{a_{UR}^{1-\sigma}}{(\sigma-\mu)(2-a_{UR}^{\gamma_m})}(1-\phi)\theta_m\left(\frac{s}{\Delta(a_{UR}, a_{LR})} - \frac{(1-s)}{\Delta^*(a_{UR}, a_{LR})}\right) - \chi, \end{aligned} \tag{15}$$

$$\begin{aligned} v_L(a_{UR}, a_{LR}) &= \pi(a_{LR}) - \pi^*(a_{LR}) - \chi \\ &= \frac{a_{LR}^{1-\sigma}}{(\sigma-\mu)(2-a_{LR}^{\gamma_m})}(1-\phi)\theta_m\left(\frac{s}{\Delta(a_{UR}, a_{LR})} - \frac{(1-s)}{\Delta^*(a_{UR}, a_{LR})}\right) - \chi \end{aligned} \tag{16}$$

where

$$\begin{aligned} \Delta(a_{UR}, a_{LR}) &= \int_0^{a_{LR}} a^{1-\sigma} dF(a) + \int_{a_{UR}}^{1} a^{1-\sigma} dF(a) + \\ & (1-s)\phi \int_{a_{LR}}^{a_{UR}} a^{1-\sigma} dF(a) + s \int_{a_{LR}}^{a_{UR}} a^{1-\sigma} dF(a), \end{aligned} \tag{17}$$

[9] Profit maximisation ensures that capital is located where its return is maximised.

[10] A much simpler, but less satisfactory approach would be to let all firms make a one shot moving decision based on the short-run return to the capital differential curve.

$$\Delta^*(a_{UR}, a_{LR}) = \phi \int_0^{a_{LR}} a^{1-\sigma} dF(a) + \phi \int_{a_{UR}}^{1} a^{1-\sigma} dF(a) + (1-s) \int_{a_{LR}}^{a_{UR}} a^{1-\sigma} dF(a) + s\phi \int_{a_{LR}}^{a_{UR}} a^{1-\sigma} dF(a) \quad (18)$$

There is no discounting. The cost of moving is given by:

$$\kappa \frac{dF(a_{LR})}{dt} + \kappa \frac{dF(a_{UR})}{dt} + \chi = \kappa \rho a_{LR}^{\rho-1}(1-s)\dot{a}_{LR} + \kappa \rho a_{UR}^{\rho-1}(1-s)\dot{a}_{UR} + \chi. \quad (19)$$

In addition to the fixed moving cost χ, the cost of moving from the periphery to the core will also depend on a congestion cost. That is, the more firms that move at the same time, the higher the migration costs. Firms from the high and low end of the distribution of a contribute to the congestion at the border. However, at the outset, firms from the low end of the distribution of a will start since the return to capital differential in (12) goes to infinity for $a = 0$. As the pressure to migrate declines, firms with a higher a start to move and at the point where the return to capital differential for the marginal mover from the low end of the distribution is the same as for the firm with $a = 1$, migrations start from both ends. Figure 3 schematically shows the dynamics.

The resulting long-run equilibrium distribution of firms is illustrated by Figure 4

Note also that since $\pi(a) - \pi^*(a) = \frac{a^{1-\sigma}}{(\sigma-\mu)(2-a^{\gamma_m})}\mu\theta_m(1-\phi)\left(\frac{s}{\Delta(a_{UR},a_{LR})} - \frac{(1-s)}{\Delta^*(a_{UR},a_{LR})}\right) = \chi > 0$ for a_U and a_L in the long-run equilibrium, there is never any tendency for movement from the large to the small region, meaning that we have one-way sorting.

The long-run a_L is determined by the condition that $v_L(a_{UR}, a_{LR}) = 0$, while a_U is given by $v_U(a_{UR}, a_{LR}) = 0 \; if \; a_U \in (0, 1)$, otherwise $a_U = 1$. The latter condition incorporates the fact that it is the firms from the low end that start moving, and firms from the other end of the distribution only start moving when trade freeness has reached $\phi = \widehat{\phi}$. Using $a_U = 1$ in (12) gives $\widehat{\phi} = 1 - \frac{\chi(\sigma-\mu)}{\mu\theta_m B}$ which is exogenous from the point of view of the firm.

Specifically, the long-run equilibrium cutoffs, a_U, a_L, are solved by

$$\begin{aligned} v_U(a_U, a_L) &= \pi(a_U) - \pi^*(a_U) - \chi \\ &= \frac{a_U^{1-\sigma}}{(\sigma-\mu)(2-a_U^{\gamma_j})}(1-\phi)\theta_m\left(\frac{s}{\Delta(a_U, a_L)} - \frac{1-s}{\Delta^*(a_U, a_L)}\right) - \chi = 0 \quad (20) \\ v_L(a_U, a_L) &= \pi(a_L) - \pi^*(a_L) - \chi \\ &= \frac{a_L^{1-\sigma}}{(\sigma-\mu)(2-a_L^{\gamma_j})}(1-\phi)\theta_m\left(\frac{s}{\Delta(a_U, a_L)} - \frac{1-s}{\Delta^*(a_U, a_L)}\right) - \chi = 0 \quad (21) \end{aligned}$$

where

$$\begin{aligned} \Delta &= \lambda[a_L^{1-\sigma+\rho} + 1 - a_U^{1-\sigma+\rho} + (1-s)\phi(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho}) + s(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho})], \quad (22) \\ \Delta^* &= \lambda[\phi(a_L^{1-\sigma+\rho} + 1 - a_U^{1-\sigma+\rho}) + (1-s)(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho}) + s\phi(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho})] (23) \end{aligned}$$

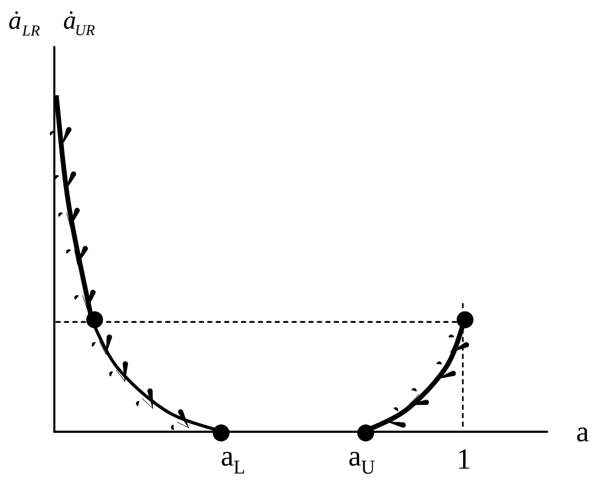

Figure 3: The dynamics of the model

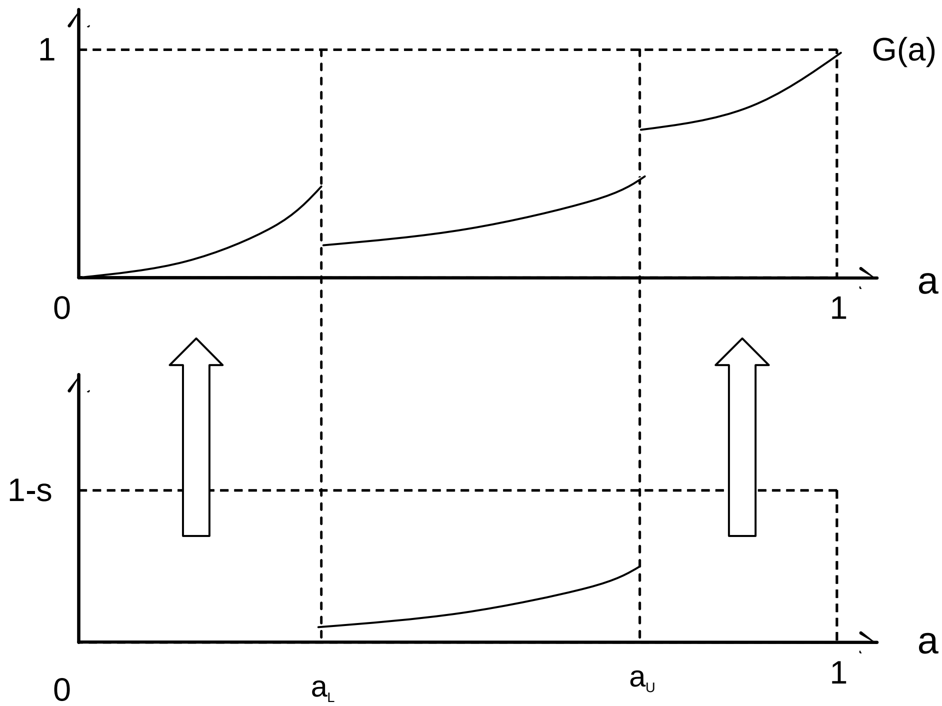

Figure 4: The long-run distribution of firms

where $\lambda \equiv \frac{\rho}{1-\sigma+\rho} > 1$. Here, we assume that $1-\sigma+\rho > 0$ to ensure convergence of the integrals. Since $\rho > 1$, this condition implies that $\rho > \sigma - 1 > 1$.

The cutoff relations: Unless the equilibrium is a corner solution, i.e. $a_U \in (0,1)$ and $a_L \in (0,1)$ ($a_L < a_U$), then $v_L(a_U, a_L) = 0$ and $v_U(a_U, a_L) = 0$ must hold simultaneously, which means that $\frac{a_U^{1-\sigma}}{(2-a_U^\gamma)} = \frac{a_L^{1-\sigma}}{(2-a_L^\gamma)}$ is always satisfied. Note that if $\gamma = 0$, we have $a_U = a_L$, which means that we have one-sided sorting as in Baldwin and Okubo (2006).

When trade costs are sufficiently high, only low a firms relocate, implying a single cut-off a_L. At a level of trade costs, ϕ^B, both low and high a firms start to migrate, leading to two-way sorting. The a_L-cutoff when the lowest productivity firm ($a_U = 1$) starts to move from the small country to the large country, which is denoted as $\widehat{a}_L$, is given by the solution of $\widehat{a}_L^{1-\sigma} - 2 + \widehat{a}_L^\gamma = 0$, and the sustainpoint ϕ^S , which is the trade costs with $a^S = a_U = a_L$, is defined by

$$\frac{(a^S)^{1-\sigma}}{(\sigma-\mu)(2-(a^S)^{\gamma_m})\lambda}(1-\phi^S)\theta_m\left(s - \frac{1-s}{\phi^S}\right) = \chi. \tag{24}$$

Closed form solutions for the long-run critical values of a are hard to obtain. Therefore we simulate the model. Figure 5 simulates the effect of reducing χ on a_U and a_L for some typical parameter values ($\mu = 0.3, \sigma = 2, \rho = 3, \gamma = 2, \phi = 0.7, s = 0.6$). Note how only high-productive firms (firms with a low a) move at the beginning. As χ reaches a sufficiently low level also firms with low productivity from the other end of the distribution start to move, thus leading to two-sided sorting.

The effects of trade integration on the long-run values of a_U and a_L are displayed in Figure

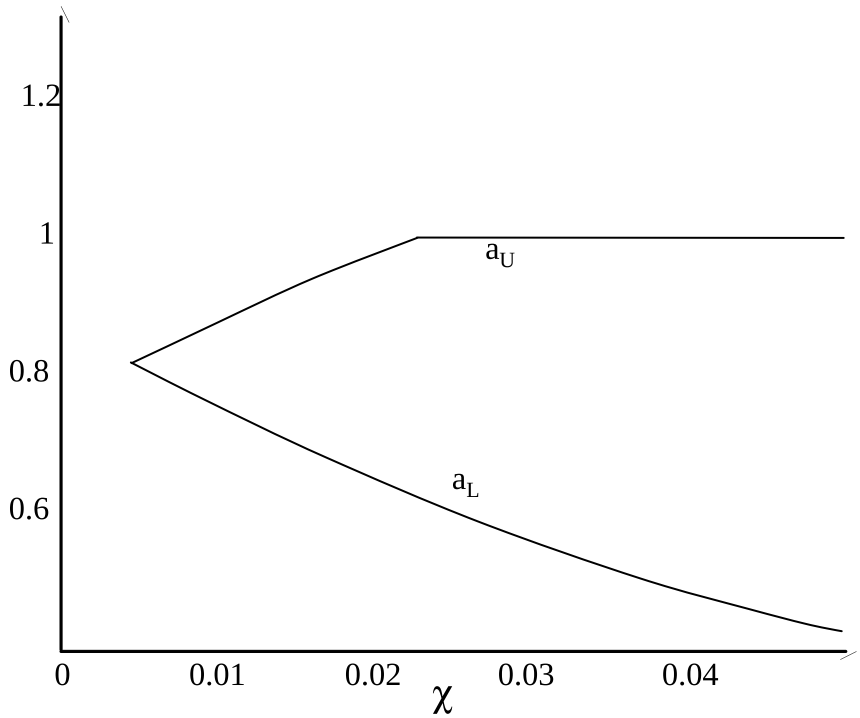

Figure 5: The effect of lower moving cost

6 for the same parameter valuse. Stronger agglomeration forces imply that the relative return to capital increases in the large region (Core). This means that the U-shaped curve in Figure 2 shifts upwards, leading to convergence of a_U and a_L. However, agglomeration forces are U-shaped in ϕ in this type of model, and Figure 6 therefore shows how a_U and a_L first converge and thereafter diverge as trade costs are reduced. Agglomeration forces are maximal at the point where the distance between a_U and a_L is smallest.

Finally, maintaining the same parameter values, Figure 7 illustrates the effect of γ. Sectors with a higher γ tend to have more two-sided sorting. However, the sorting from the low end is U-shaped in γ.

2.4 Average Productivity

A key feature of the Melitz model is that productivity increases due to trade liberalisation as the least productive firms disappear. Here no firms die, since there are no entry costs, but firms move and this affects sector productivity and therefore, average productivity. Productivity in a sector in the two economies can be defined as a frequency weighted mean of individual productivities (see Melitz 2003):

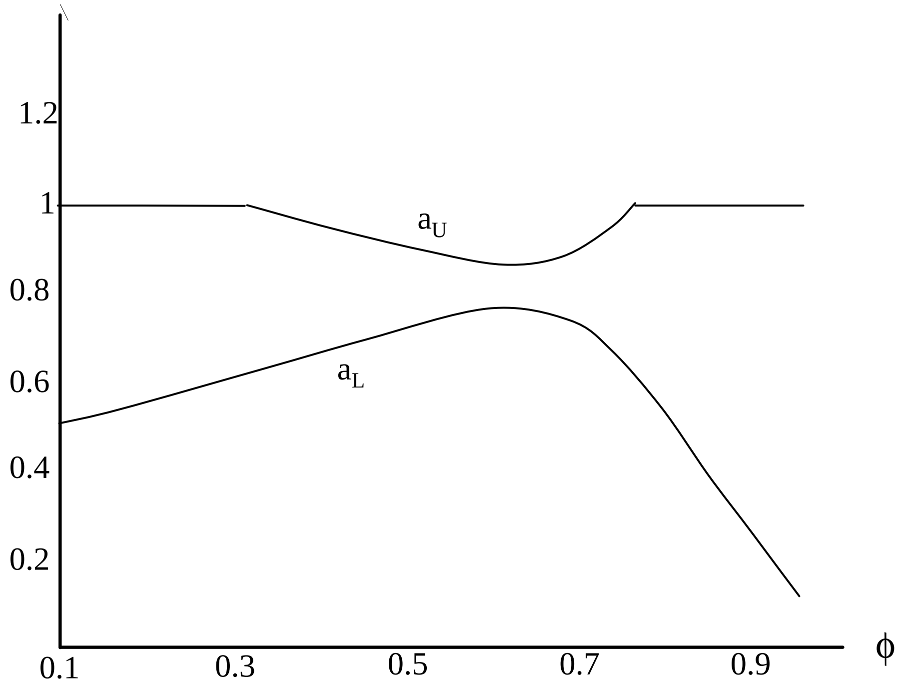

Figure 6: The effect of lower trade costs

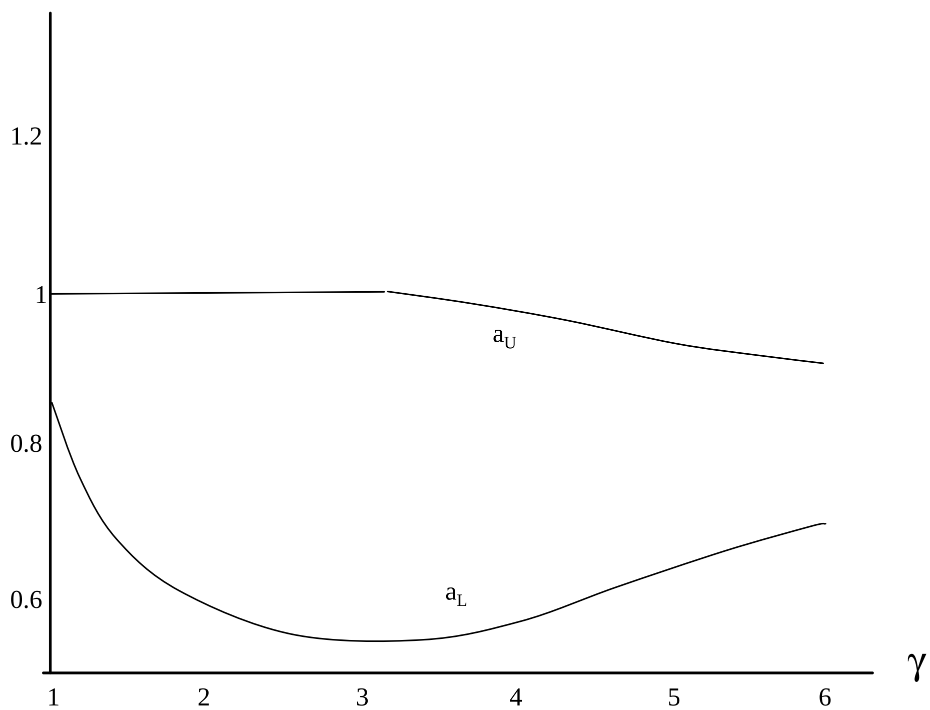

Figure 7: The effect of γ

$$\begin{aligned}
\varphi &= \left(\int_0^{a_L} a_i^{1-\sigma} dF(a) + \int_{a_U}^{1} a_i^{1-\sigma} F(a) + s \int_{a_L}^{a_U} a_i^{1-\sigma} dF(a) \right)^{\frac{1}{\sigma-1}} \\
&= \left(\lambda - \lambda(1-s)(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho}) \right)^{\frac{1}{\sigma-1}}, \\
\varphi^* &= \left((1-s) \int_{a_L}^{a_U} a_i^{1-\sigma} dF(a) \right)^{\frac{1}{\sigma-1}} \\
&= \left(\lambda(1-s)(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho}) \right)^{\frac{1}{\sigma-1}}. \qquad (25)
\end{aligned}$$

We show in the appendix that $\frac{\partial(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho})}{\partial\chi} > 0$, $\frac{d(a_U - a_L)}{d\gamma} < 0$, and $\frac{\partial(a_U^{1-\sigma+\rho} - a_R^{1-\sigma+\rho})}{\partial\phi} \gtrless 0$. First, a reduction in the fixed moving cost leads to a fall in average productivity in the Periphery and an increase in the Core. It is also the case that sectors with high capital requirements tend to have a higher productivity in the large region and a lower productivity in the small region. That

Moreover, in spite of two-way sorting, it is always the case that productivity is higher in the core as illustrated by using (19):

$$\varphi^{\sigma-1} - \varphi^{*\sigma-1} = \lambda(1 - 2(1-s)(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho})) > 0 \qquad (26)$$

Finally, the results for trade liberalisation are ambigous.

2.5 Two-sided Sorting

A distinctive feature of our model is the occurrence of two-sided sorting. The simulations above indicate that two-sided sorting increases as the fixed moving cost χ is reduced. This result is formally seen from the result that $\frac{\partial(a_U^{1-\sigma+\rho} - a_L^{1-\sigma+\rho})}{\partial\chi} < 0$, where $1 - \sigma + \rho > 0$.

Second, the result that $\frac{d(a_U - a_L)}{d\gamma} < 0$ implies that the degree of two-sided sorting increases in sectors with a higher average capital labour ratio.[11] We test this property in the empirical section below.

3 Empirical Analysis

3.1 Empirical Strategy

While agglomeration, selection, one-sided sorting, and two-sided sorting all lead to higher average productivity in the core region, they have different implications for the distribution of firm productivities in the core versus the periphery. Figure 8 schematically shows the four

[11] The result is derived under the condition that $a_U < 1$. That is, the result does not necessarily hold when there is only one-sided sorting, as illustrated in Figure 7.

cases. The solid line in the figure indicates distribution in Core and the dotted line indicates the distribution in Periphery. Figure 8a shows a sector with a pattern consistent with standard agglomeration models, implying that all firms benefit from being in the core, thus implying that the productivity distribution of the core firms is shifted to the right as compared to the distribution of the firms in the periphery.[12] Next, Figure 8b shows the selection case as in a standard heterogeneous-firm model where the distribution in the core is left truncated. Finally, Figures 8c and 8d illustrate sorting. In Figure 8c, which illustrates one-sided sorting, the distribution of the periphery is truncated from the right because the most productive firms migrate to the centre, thus producing an upward jump in the distribution in the core. Figure 8d shows the case modelled in this paper where the periphery is truncated from both sides as a result of two-sided sorting.

[12] This is the cleanest case. Naturally, it is possible to assume e.g. that more productive firms have better capacity to absorb spillovers, in which case the shape of the distribution will be affected in addition to the shift.

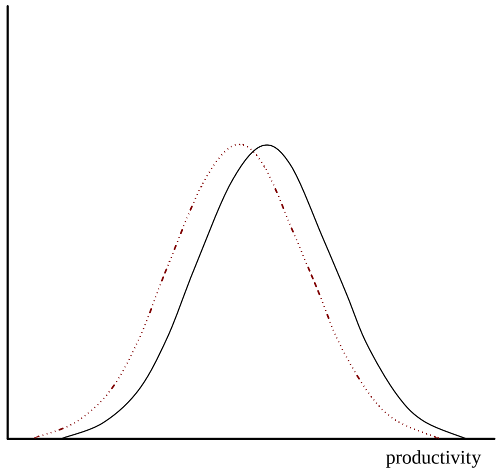

Figure 8a: Agglomeration

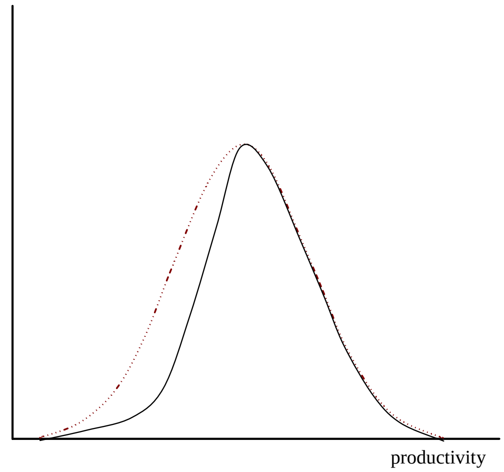

Figure 8b: Selection

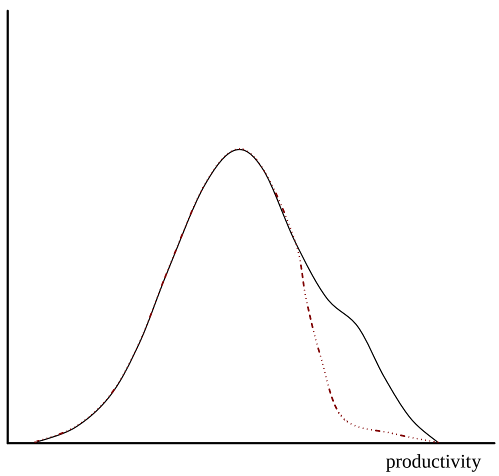

Figure 8c: One-sided sorting

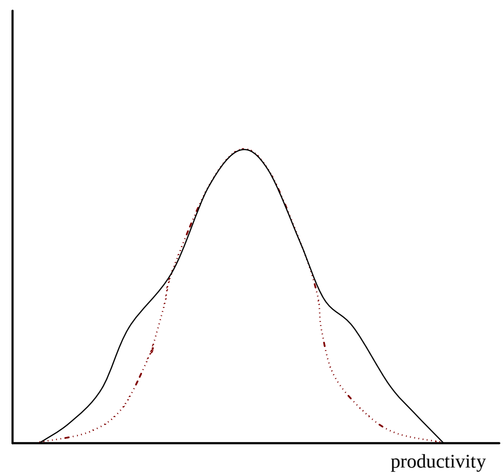

Figure 8d: Two-sided sorting

These patterns imply a number of testable hypotheses concerning differences between the core and the periphery of all first three moments of the productivity distribution of firms. Table 1 shows the predictions for the moments in the four cases shown in Figure 8, where superscript "c" indicates the core. Pure agglomeration only affects the mean ($\bar{x}$). Selection and sorting affect all three moments but the important difference is a different sign of the spread-gap, $s^c - s$, between the core and the periphery. Selection reduces the spread of the distribution of the core, whereas sorting instead leads to a larger spread in the core. Finally, the difference between one- and two-sided sorting is just that the difference in skewness (g) between the core and the periphery is smaller in the case of two-sided sorting.

	mean	spread	skewness
agglomeration	$\overline{x}^c > \overline{x}$	$s^c = s$	$g^c = g$
selection	$\overline{x}^c > \overline{x}$	$s^c < s$	$g^c < g$
one-sided sorting	$\overline{x}^c > \overline{x}$	$s^c > s$	$g^c < g$
two-sided sorting	$\overline{x}^c > \overline{x}$	$s^c > s$	$g^c \approx g$

(Table 1)

Clearly, several of these mechanisms may be active simultaneously in practice. The question is therefore rather which of them dominate. The answer, as we will see next, differs depending on which sector we study.

3.2 Data

Here we use firm (plant) level data from Japan's *Census of Manufacturers* (METI) virtually covering all plants with more than five employees in 1990, classified at the three-digit sector level.[13]. In total, 324,000 plants and 154 sectors. The sector classification is shown in the appendix. The manufacturing census contains basic information on plants, such as output (shipment) and employment (number of regular workers), but no identifier linking firms under the same ownership. Hence, aggregation of the data is not possible.

There are 47 prefectures[14]. We define the core region as the 16 central prefectures surrounding Tokyo, Osaka(the second largest), Nagoya (the third) as well as Fukuoka (the fourth) prefecture. Together, they constitute the Japanese manufacturing belt. The peripheral regions are defined as the other 30 prefectures in the mainland (excluding Okinawa).

Productivity is measured by value added (unit: million yen) per regular number of employees. The capital labour ratio is measured by capital asset (unit: million yen) per employed individuals. All variables are in logs. Descriptive statistics are shown in the appendix. The regional GDP measure is taken from Fukao and Yue (2000).

3.3 Analysis

First, Table 2 shows that our data at the sectoral level has standard properties. The productivity gap between the core and the periphery, measured as the difference in value added per employee (in logs) for plants in the central districts of Japan compared to plants in peripheral districts at the three-digit sectoral level, increases with the distance between the core and the periphery and decreases with the size of the periphery. As a proxy for trade cost, we use the minimum geographical distance from the bipolar largest cities (Tokyo or Osaka) for 46 prefec-

[13] 1990 is the last period of interregional relocation within Japan. From the mid 1990s and onwards, Japanese firms became very active in FDI and outsourcing, which may blur the pattern of interregional relocation within Japan.

[14] The Japanese prefectures are administrative units similar to the NUTS2 regions in EU. The Okinawa island is excluded. Thus our data sample is 46 prefectures.

Table 2: Core region productivity premium

Productivity Gap	
logDist	0.049***
	(3.79)
logGDPperif	-0.066***
	(-4.01)
Const.	0.98***
	(3.52)
R2-adj.	0.01
F-stat	19.42
N.obs.	3481

t-statistic in parenthesis. *=10%,**=5%, and ***=1% significance level.

tures (excluding Okinawa) (unit: km). As shown by Okubo and Tomimira (2010a), similar properties hold for this dataset when the data is aggregated.

Our variables of interest are sectoral differences (gaps) in mean productivity, standard deviation, and skewness between the core and the periphery: $\bar{x}^c - \bar{x}$, $s^c - s$, and $g^c - g$. The theoretical models discussed above predict a higher productivity in the core and we therefore focus on sectors where $\bar{x}^c - \bar{x} > 0$. Among these, sectors with a negative spread gap $s^c - s < 0$ are considered to be dominated by selection, and sectors with $s^c - s > 0$ to be dominated by sorting (compare Table 1). Finally, we make a difference between one- and two-sided sorting by looking at the skewness gap, $g^c - g$. We will label sectors as subject to two-way sorting when the skewness gap is not too large. For illustrative purposes, we label a sector as subject to two-sided sorting when $|g^c - g| < 0.5$. Naturally, the exact limit between one- and two-sided sorting is arbitrary.

Figures 9a-d show a few examples of our estimated kernel density function in core and periphery, respectively. The figures single out representative sectors that are classified as agglomeration (Figure 9a), selection (Figure 9b), one-sided sorting (Figure 9c) and two-sided sorting (Figure 9d). As illustrated by the figures, real world cases are less clear than the stylised theoretical cases, and several of the above mentioned mechanisms behind the higher productivity in the core could certainly be present at the same time. The question is therefore rather which mechanism tends to dominate for each sector.

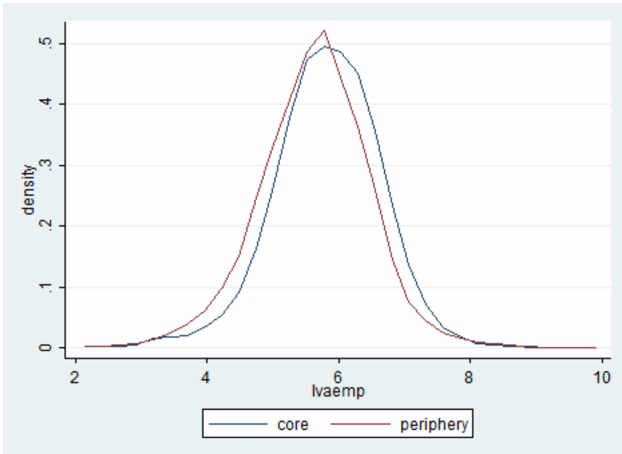

Figure 9a: Processing of Fish and Fish Products

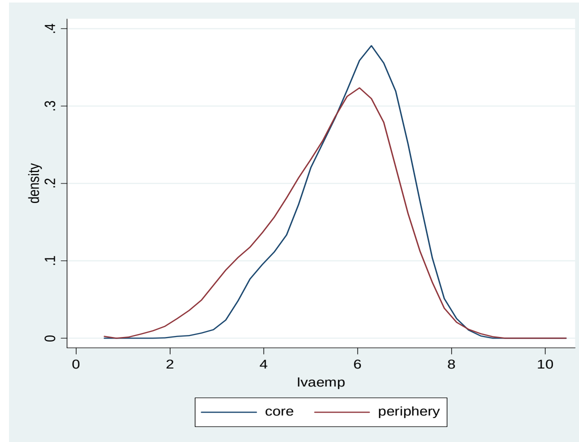

Figure 9b: 133 Tea and Coffe

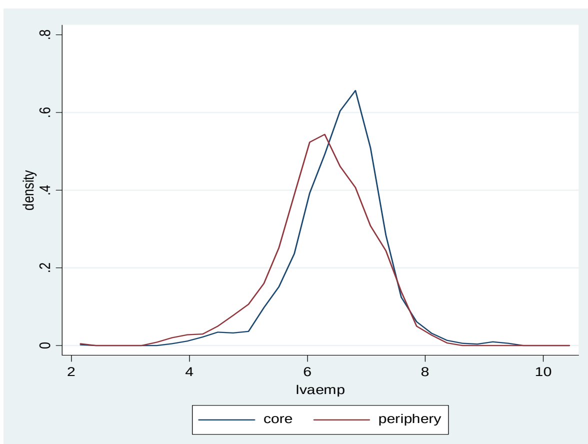

Figure 9c: 205 Oil and fat products

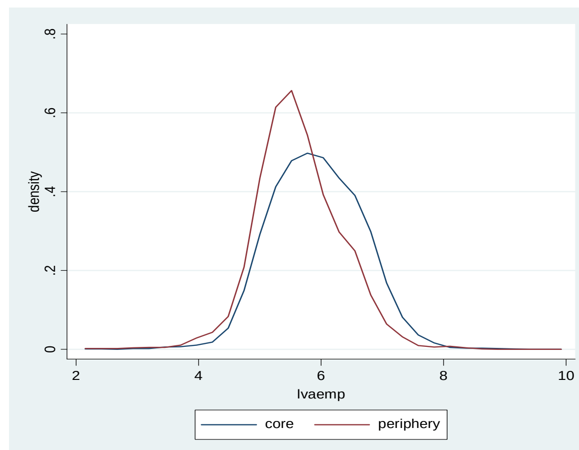

Figure 9d: 308 Parts and components to electronic devices

A comprehensive picture is given by Figure 10 which plots the spread gap, $s^c - s$, against the skewness gap, $g^c - g$, for all sectors in our sample with a higher average productivity in the core, $\bar{x}^c - \bar{x} > 0$, (numerical values for all sectors are shown in Table A2 in the appendix). Each dot indicates the spread and skewness gaps in each of the 30 peripheral prefectures against the average of the 16 core prefectures. Using the classification in Table 1, sectors in the South Western quadrant would be classified as dominated by selection, whereas the North Western quadrant are sectors dominated by sorting. Two-sided sorting would produce a smaller skewness-gap and they are therefore located closer to the vertical zero skewness-gap line, whereas sectors dominated by one-sided sorting would lie further to the left in the North-Western quadrant. The general picture is that both sorting and selection seem to be present in a large number of cases.[15]

[15] It is possible that agglomeration externalities would be biased e.g. so that more productive firms have a better absorbtive capacity for positive spillovers as in Combes et al. (2009). This would generate positive skewness and

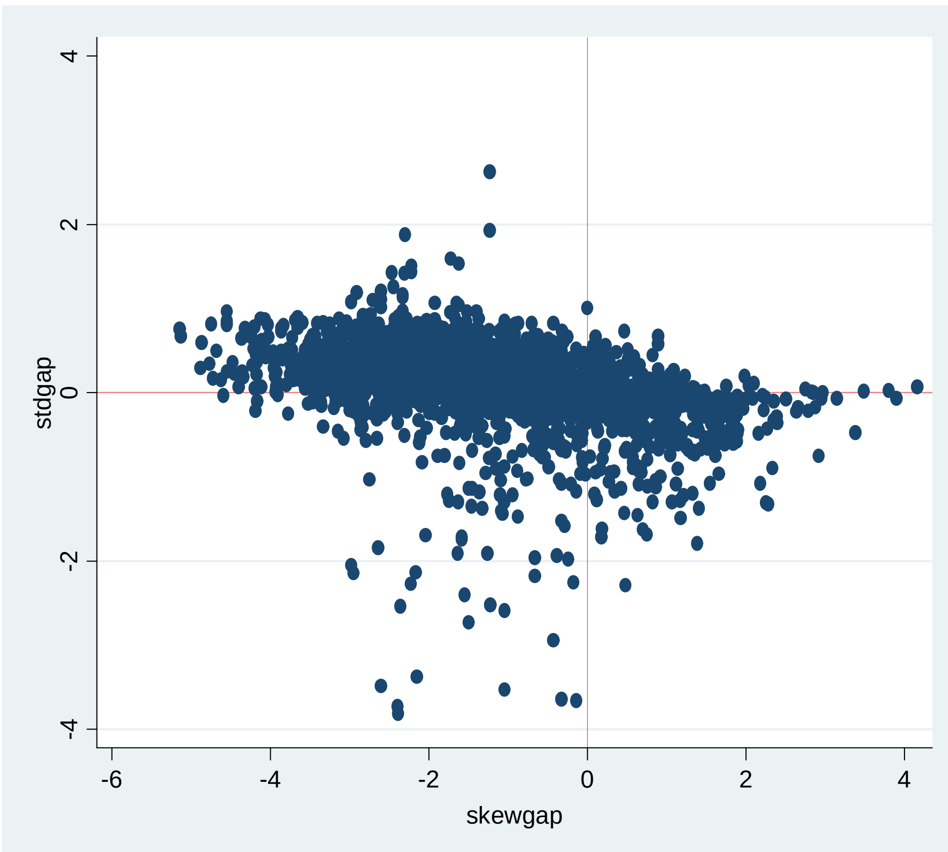

Figure 10: Sorting and selection

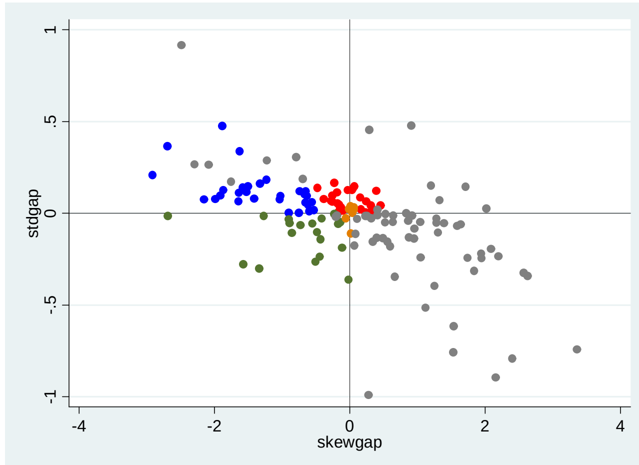

Figure 11: Aggregated sorting and selection

To get a clearer picture, Figure 11 plots the spread gap against the skewness gap with all peripheral regions aggregated. Grey indicates sectors that do not fit the classification in Table 1 ($\bar{x}^{c} - \bar{x}^{-} < 0$). The aggregated figure makes it easier to illustrate the four different cases of interest: Green indicates selection and yellow agglomeration. Red and blue are two-sided and one-sided sorting. A relatively large number of sectors located in the South-Western quadrant have a pattern consistent with selection (green), and a similar number of sectors consistent with sorting (blue and red). Fewer sector could be classified as pure agglomeration (yellow). The figure is illustrative and the classification of agglomeration and two-sided and one-sided sorting based on the size of the skewness gap is naturally arbitrary. Choosing a more narrow definition of two-sided sorting would e.g. shrink the cluster of red points from both sides and expand the number of blue points. Likewise would a tighter limit of skewness gap for agglomeration shrink the yellow point cluster.

3.4 Capital intensity and sorting

Our model associates two-sided sorting with high capital intensity of a sector as illustrated in Figure 7. Two-sided sorting reduces the skewness in peripheral regions, and we would therefore expect to see a negative relationship between capital intensity and the skewness gap. Table 3 shows that the skewnessgap is robustly negatively related to the capital labour ratio of a sector. Also the effect of distance is estimated with the expected sign since longer distance implies higher trade costs and therefore less sorting. The realtionship is also robust to the inclusion of several contrpls such as the size of the sector (measured by employment):

We now turn to investigating skewness in the periphery, which may be an even more direct measure of two-sided sorting. Table 4 shows how the skewness in peripheral regions at the sectorial level is negatively associated with the capital labour ratio. This relationship is once

higher spread in the core corresponding to the Northeastern quadrant in the figure.

Table 3 The sectorial skewness gap

Skewness-Gap	1	2	3
logK/L	-0.19***	-0.20***	-0.19***
	(-5.77)	(-5.89)	(-5.73)
logDist		-0.11***	-0.10***
		(-3.30)	(-3.02)
logemp			-0.088*
			(-1.66)
Const.	-0.63***	-0.021	0.13
	(-3.44)	(-0.09)	((0.55)
R2-adj.	0.012	0.016	0.016
F-stat	33.32	22.16	15.7
N.obs.	2644	2644	2644

t-statistic in parenthesis. *=10%,**=5%, and ***=1% significance level.

more robust to the inclusion of several control variables such as distance and firms' employment level (firm size).

The distinct feature of two-sided sorting in our model, compared to e.g. one-sided sorting or selection, is that firms with a low productivity move from the periphery to the core. The degree to which this will happen depends on the capital intensity of a sector. To measure this effect as directly as possible, we calculate the productivity level by sector in the periphery for which the cumulative density is 25 percent, and relate this measure to the capital intensity of the sector. The productivity distribution starts at zero in each sector.[16] Sorting from the low end means that the productitvty distribution is hollowed out at the low end, which means that the 25 percentile productivity level becomes higher. Thus, the model predicts a positive relationship between the lower 25 percentile productivity level and the capital intensity of a sector in the periphery.

Figure 12 plots this relationship for all sectors. There seems to be a robust positive relationship for a large group of sectors but also a very different pattern for a large group of outliers. Regressing the sectorial capital labour ratio on the lower 25 percentile productivity level does not produce a significant positive relationship. The outliers are primarily sectors with few large and badly performing firms implying that most of the mass of the productivty distribution is concentred close to zero.[17] Requiring the standard deviation to be larger than 0.7 weeds out some of the sectors with the most concentrated productivity distribution. Table 5 shows the regression results for this sample. The positive relationship is robust to controlling for the size

[16] Some firms display a negative value added per employee. We have set these to zero in order to be able to take logs.

[17] Firms with a negative value added are set to zero in order to be able to take logs.

Table 4 Skewness in peripheral regions by sector

Skewness in periphery	1	2	3
logK/L	-0.065**	-0.063**	-0.075***
	(-2,27)	(-2.18)	(-2.62)
logDist		0.079***	0.057**
		(2.73)	(1.96)
logemp			0.24***
			(5.15)
Const.	-0.37***	-0.81***	-1.22***
	(-3.44)	(-4.18)	(-5.85)
R2-adj.	0.0016	0.004	0.014
F-stat	5.15	6.3	13.07
N.obs.	2644	2644	2644

t-statistic in parenthesis. *=10%,**=5%, and ***=1% significance level.

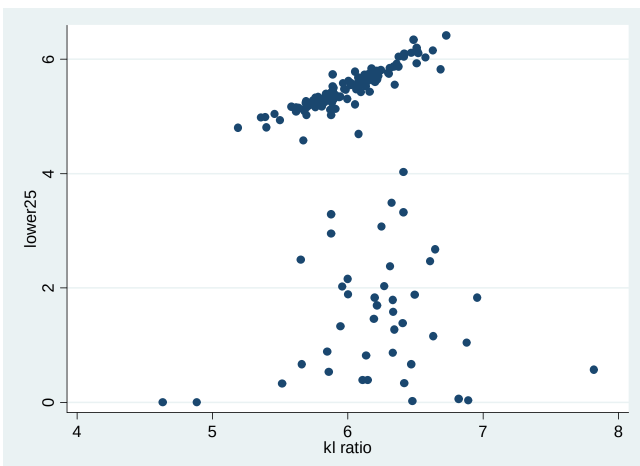

Figure 12: Productivity against the the capital labour ratio for the 25 percent lower tail of the productivity distribution for different sectors.

Table 5 The productivity level at the 25 percent percentile

Productivity at the 25 percent percentile	1	2
logK/L	0.86**	1.1***
	(2.12)	(2.79)
logemp		-1.30***
		(-3.39)
Const.	-0.43	1.13
	(-0.18)	(0.47)
R2-adj.	0.028	0.11
F-stat	4.48	8.19
N.obs.	120	120

t-statistic in parenthesis. *=10%,**=5%, and ***=1% significance level.

of the sector in terms of employment. Distance does not apply since we are just regressing sectors in the periphery.

4 Conclusion

This paper develops a model of two-sided spatial sorting, where high-productivity firms with a high capital intensity and low-productivity firms with a low capital intensity tend to locate in the large core region. Firms with intermediate productivity and capital intensity remain in the periphery. We show that a reduction in the fixed moving cost leads to a fall in average productivity in the small foreign country and an increase in the large home economy. It is also the case that sectors with high capital requirements will have a higher productivity in the large region than in the small region.

To empirically distinguish between standard agglomeration externalities, sorting and selection we note that sorting has a very different implication for the second- and third-order moments of the productivity distribution of firms in the different regions. While externalities have no effect on the distribution spread and selection reduces the spread in the core, sorting will increase the spread in the core, as firms from possibly both ends of the productivity distribution move from the periphery to the core. In the case of two-sided sorting, the effect on the skewness of the distribution in the core and the periphery may be weaker than in the case of one-sided sorting, in which case only firms from the upper tail move.

We use data from Japan's *Census of Manufacturers* covering virtually all plants with more than five employees in 1990, classified at the three-digit sector level, to investigate the predictions of the model. A problem here is that agglomeration, selection and sorting can all be present at the same time, and the question is therefore rather which of these forces dominate for a specific sector. When plotting the difference in the distribution spread between the core and the periphery against the difference of skewness between these, selection and sorting seem to be dominating for a roughly equal share of the sectors.

A main result from our model is that the tendency for two-sided sorting is positively related to the capital labour ratio of a sector. We test this prediction by comparing the skewness of the core and the peripheral regions, but also by directly analysing the peripheral distributions. Empirical evidence supports the predicition that sectors with a high capital labour ratio have relatively less low-productivity firms in peripheral regions.

5 Appendix

5.1 Proofs

5.1.1 Proof that $\frac{d(a_U - a_L)}{d\chi} > 0$

Using the total differentiation of $v_U(a_U, a_L) = 0$,

$$\frac{dv_U}{da_U} = \frac{d\Omega}{da_U}(1-\phi)B + \Omega(1-\phi)\left(\frac{dB}{d\Delta}\frac{d\Delta}{da_U} + \frac{dB}{d\Delta^*}\frac{d\Delta^*}{da_U}\right),$$

$$\frac{dv_U}{da_L} = \frac{d\Omega}{da_L}(1-\phi)B + \Omega(1-\phi)\left(\frac{dB}{d\Delta}\frac{d\Delta}{da_L} + \frac{dB}{d\Delta^*}\frac{d\Delta^*}{da_L}\right).$$

where $\Omega \equiv \frac{a^{1-\sigma}}{(\sigma-\mu)(2-a^\gamma)}$. Since $\frac{d\Delta}{da_U} = -\frac{d\Delta}{da_L}$, and $\frac{d\Delta^*}{da_U} = -\frac{d\Delta^*}{da_L}$

$$\frac{dv_U}{da_U} - \frac{dv_U}{da_L} = \left(\frac{d\Omega}{da_U} - \frac{d\Omega}{da_L}\right)(1-\phi)B + 2\Omega(1-\phi)\left(\frac{dB}{d\Delta}\frac{d\Delta}{da_U} + \frac{dB}{d\Delta^*}\frac{d\Delta^*}{da_U}\right).$$

Since $\frac{dB}{d\Delta} < 0$, $\frac{d\Delta}{da_U} < 0$, $\frac{dB}{d\Delta^*} > 0$, $\frac{d\Delta^*}{da_U} > 0$ the sign of this expression depends on the first term $\left(\frac{d\Omega}{da_U} - \frac{d\Omega}{da_L}\right)$. Since $a_U > a_L$ this depends in the signs of $\frac{d^2\Omega}{da^2}$. First

$$\frac{d\Omega}{da} = \frac{(1-\sigma)a^{-\sigma}(2-a^\gamma) + \gamma a^{\gamma-\sigma}}{(\sigma-\mu)(2-a^\gamma)^2},$$

The denominator is decreasing in a. Differentiating the numerator w.r.t. a gives

$$\begin{aligned} a^{-\sigma-1}\left((2\sigma(\sigma-1) + (\sigma-\gamma)(1-\gamma-\sigma)a^\gamma\right) &> a^{-\sigma-1}(2\sigma(\sigma-1) + (\sigma-\gamma)(1-\gamma-\sigma)) \\ &= a^{-\sigma-1}(\sigma-1)(\sigma-1+\gamma) \end{aligned}$$

the righ hand side expression is positive for $\sigma - 1 + \gamma > 0$, which is assured by our previous assumptions that $1 - \sigma + \rho > 0$, and the condition that $\rho > 1$. Therefore $\frac{d^2\Omega}{da^2} > 0$ which implies that $\left(\frac{d\Omega}{da_U} - \frac{d\Omega}{da_L}\right) > 0$. We therefore have that

$$\frac{dv_U}{da_U} - \frac{dv_U}{da_L} > 0.$$

Finally, it is easy to see that

$$\frac{dv_U}{d\chi} < 0$$

As a result of these differentiations,

$$\frac{d(a_U - a_L)}{d\chi} = -\frac{d(a_U - a_L)/dv_U}{d\chi/dv_U} > 0$$

5.1.2 Proof that $\frac{d(a_U - a_L)}{d\phi} \lesseqgtr 0$

We can derive

$$\frac{dv_U}{d\phi} = -\Omega B + \Omega(1-\phi)\left(\frac{dB}{d\Delta}\frac{d\Delta}{d\phi} + \frac{dB}{d\Delta^*}\frac{d\Delta^*}{d\phi}\right) \lesseqgtr 0,$$

,where $\Omega \equiv \frac{a^{1-\sigma}}{(\sigma-\mu)(2-a^{\gamma})}$. The first term is negative but the second is positive due to $\frac{dB}{d\Delta^*}\frac{d\Delta^*}{d\phi} > \frac{dB}{d\Delta}\frac{d\Delta}{d\phi}$.

Thus,

$$\frac{d(a_U - a_L)}{d\phi} = -\frac{d(a_U - a_L)/dv_U}{d\phi/dv_U} \lesseqgtr 0$$

5.1.3 Proof that $\frac{d(a_U - a_L)}{d\gamma} < 0$

$$\frac{dv_U}{d\gamma} = (1-\phi)B\frac{d\Omega}{d\gamma} > 0,$$

,where $\Omega \equiv \frac{a^{1-\sigma}}{(\sigma-\mu)(2-a^{\gamma})}$. This follows from the fact that

$$\frac{d\ln\Omega}{d\gamma} = -\frac{1}{2-a^{\gamma}}a^{\gamma}\ln a > 0$$

Thus,

$$\frac{d(a_U - a_L)}{d\gamma} = -\frac{d(a_U - a_L)/dv_U}{d\gamma/dv_U} < 0$$

Furthermore, using $\frac{dv_U}{d\gamma} > 0$ and $\frac{dv_U}{d\chi} < 0$, we can show that $\frac{d\chi}{d\gamma} = -\frac{dv_U/d\gamma}{dv_U/d\chi} < 0$.

5.2 Sector Classification: Table A1

sector		sector	
121	Livestock products	243	Cut stock and findings for boots and shoes
122	Seafood products	244	Leather footwear
123	Canned and preserved fruit and vegetable products	245	Leather gloves and mittens
124	Seasonings	246	Luggage
125	Sugar processing	247	Handbags and small leather cases
126	Flour and grain mill products	248	Fur skins
127	Bakery and confectionery products	249	Miscellaneous leather products
128	Animal and vegetable oils and fats	251	Glass and its products
129	Miscellaneous foods and related products	252	Cement and its products
131	Soft drinks and carbonated water	253	Structural clay products, except those of pottery
132	Alcoholic beverages	254	Pottery and related products
133	Tea and coffee	255	Clay refractories
134	Manufactured ice	256	Carbon and graphite products
135	Prepared animal foods and organic fertilizers	257	Abrasive products
141	Silk reeling plants	258	Aggregate and stone products
142	Spinning mills	259	Miscellaneous ceramic, stone and clay products
143	Twisting and bulky yarns	262	Iron smelting, without blast furnaces
144	Woven fabric mills	263	Steel, with rolling facilities
145	Manufacturing kni	264	Steel materials, except made by smelting furnaces and with rolling facilities
146	Dyed and finished textiles	265	Coated steel
147	Rope and netting	266	Forging steel manufacturing forged products
148	Lace and other textile goods	267	Pig iron article of cast metal manufacturing
149	Miscellaneous textile mill products	269	Miscellaneous iron and steel
151	Textile outer garments, except japanese style	271	Primary smelting and refining of non-ferrous metals
152	Shirts and Underwear, except japanese style	272	Secondary smelting and refining of nonferrous metals
153	Hat manufacturing	273	Rolling of non-ferrous metals and alloys, including drawing and extruding
154	Fur apparel and apparel accessories	274	Non-ferrous metal machine parts and tooling products
155	Other textile apparel and accessories, including japanese style	275	Electric wire and cable
159	Miscellaneous fabricated textile products	279	Miscellaneous non-ferrous metal products
161	Sawing, planning mills and wood products	281	Tin cans and other plated sheet products
162	Millwork, plywood and prefabricated structural wood products	282	Tableware (occidental type), cutlery, hand tools and hardware
163	Wooden, bamboo and rattan containers	283	Heating apparatus and plumbing supplies
164	Wooden footwear manufacturing	284	Fabricated constructional and architectural metal products
169	Miscellaneous manufacture of wood products	285	Metal machine parts and tooling products
171	Furniture	286	Metal coating, engraving and heat reating, except enameled ironware
172	Furniture for religious purposes	287	Fabricated wire products
173	Sliding doors and screens	288	Bolts, nuts, rivets, machine screws and wood screws
179	Miscellaneous furniture and fixtures	289	Miscellaneous fabricated metal products
181	Pulp	291	Boilers, engines and turbines
182	Paper	292	Agricultural machinery and equipment
183	Coated and glazed paper	293	Machinery and equipment for construction

183	Coated and glazed paper	293	Machinery and equipment for construction and mining, including tractors
184	Paper products	294	Metal working machinery
185	Paper containers	295	Textile machinery
189	Miscellaneous pulp, paper and paper worked products	296	Special industry machinery
191	Newspaper industries	297	General industry machinery and equipment
192	Publishing industries	298	Office, service industry and household machines
193	Printing, except mimeograph printing industries	299	Miscellaneous machinery and machine parts
194	Plate making for printing	301	Electrical generating, transmission, distribution and industrial apparatus
195	Bookbinding and printed matter	302	Household electric appliances
199	Service industries related to printing trade	303	Electric bulbs and lighting fixtures
201	Chemical fertilizers	304	Communication equipment and related products
202	Industrial inorganic chemicals	305	Electronic data processing machines, computers, equipment and accessories
203	Industrial organic chemicals	306	Electronic equipment
204	Chemical fibres	307	Electric measuring instruments
205	Oil and fat products, soaps, synthetic detergents	308	Electronic parts and devices
206	Drugs and medicines	309	Miscellaneous electrical machinery equipment and supplies
209	Miscellaneous chemical and allied products	311	Motor vehicles, parts and accessories
211	Petroleum refining	312	Railroad equipment and parts
212	Lubricating oils and greases (not made in petroleum refineries)	313	Bicycles and parts
213	Coke	314	Shipbuilding and repairing, and marine engines
214	Briquettes and briquette balls	315	Aircraft and parts
215	Paving materials	319	Miscellaneous transportation equipment
219	Miscellaneous petroleum and coal products	321	Measuring instruments, analytical instruments and testing machines
221	Plastic plates, bars and rods, pipes and tubes	322	Surveying instruments
222	Plastic films, sheets, floor coverings and synthetic leather	323	Medical instruments and apparatus
223	Industrial plastic products	324	Physical and chemical instruments
224	Foamed and reinforced plastic products	325	Optical instruments and lenses
225	Compounding plastic materials, including reclaimed plastics	326	Ophthalmic goods, including frames
229	Miscellaneous plastic products	327	Watches, clocks, clockwork-operated devices and parts
231	Tires and inner tubes	331	Manufacture of ordnance and accessories
232	Rubber and plastic footwear and its findings	343	Toys and sporting goods
233	Rubber belts and hoses and mechanical rubber goods products	344	Pens, lead pencils, painting materials and stationery
239	Miscellaneous rubber products	345	Costume jewellery, costume accessories, buttons and related products
241	Leather tanning and finishing	346	Lacquer ware
242	Mechanical leather products, except gloves and mittens	348	Manufacturing industries, n.e.c
		349	Manufacturing industries, n.e.c

5.3 Gap estimates per sector: Table A2

Type: , "1": 1-sided sorting case,"2": 2-sided sorting case, "3":selection case, "4":Agglomeration case

sector	x-gap	s-gap	g-gap	Type
121	0,28	0,02	-0,53	1
122	0,24	-0,11	0,02	A
123	0,24	-0,05	-0,14	S
124	0,22	0,02	0,17	2
125	0,83	-0,40	1,25	
126	0,00	0,31	-0,79	
127	0,18	-0,02	0,25	
128	0,14	0,48	-1,89	1
129	0,26	-0,05	-0,14	S
131	0,58	-0,61	1,54	
132	0,01	0,03	-0,13	2
133	0,38	-0,03	-0,90	S
134	0,03	0,03	2,02	
135	0,10	-0,27	-0,51	S
141	-0,23	0,45	0,29	
142	0,13	0,14	1,71	
143	-0,11	-0,01	-0,19	
144	-0,07	-0,02	-0,21	
145	0,25	-0,06	1,39	
146	0,02	-0,03	-0,41	S
147	0,17	-0,01	-1,27	S
148	-0,05	0,02	0,41	
149	0,04	-0,01	0,64	
151	0,18	0,00	0,53	
152	0,25	0,07	1,33	
153	0,44	-0,05	0,52	
154	0,10	0,37	-2,70	1
155	0,14	0,16	-1,33	1
159	0,04	0,01	0,36	2
161	0,11	0,12	-0,03	2
162	0,14	-0,08	0,96	
163	0,29	-0,18	0,07	
164	-0,06	-0,79	2,40	
169	0,30	-0,01	0,23	
171	0,12	0,00	0,83	
172	0,10	0,09	-1,02	1
173	0,15	0,12	-0,65	1
179	0,16	0,04	0,01	A
181	-0,28	-0,52	1,12	
182	0,73	-0,10	-0,48	S
183	0,19	-0,01	-2,69	S
184	0,07	-0,06	-0,89	S
185	0,22	-0,18	0,60	
189	0,22	-0,19	-0,12	S
191	0,23	0,13	-1,87	1
192	0,60	0,13	0,04	2
193	0,31	0,10	-0,26	2
194	0,31	0,00	0,04	A
195	0,27	-0,03	1,29	
199	0,54	0,34	-1,63	1
201	-0,43	0,48	0,91	
202	-0,08	0,26	-2,09	
203	0,19	-0,99	0,28	
204	-0,47	1,31	-2,11	
205	0,23	0,11	-1,64	1
206	0,23	-0,35	0,67	
209	0,07	-0,24	1,74	
211	1,97	-1,13	-1,61	S
212	0,28	-0,06	-0,55	S
215	0,27	0,12	-0,74	1
219	0,70	-0,90	2,16	
221	0,22	-0,23	2,20	
222	0,09	-0,16	0,56	
223	0,06	0,06	0,24	2
224	0,09	0,07	-1,03	1
225	0,23	0,10	-0,67	1
229	0,07	-0,04	0,86	
231	-0,33	0,92	-2,49	
232	0,14	-0,05	0,64	
233	0,13	-0,14	-0,43	S
239	0,28	-0,24	1,05	
241	0,30	0,18	-1,23	1
242	0,36	-0,30	-1,34	S
243	0,47	0,15	1,20	
244	0,34	-0,11	-0,86	S

Type: , "1": 1-sided sorting case,"2": 2-sided sorting case, "3":selection case, "4":Agglomeration case

sector	x-gap	s-gap	g-gap	Type	sector	x-gap	s-gap	g-gap	Type
245	0,40	-0,76	1,53		294	0,09	0,01	0,32	2
246	0,23	-0,31	1,84		295	0,04	0,06	-1,64	1
247	0,21	0,08	0,16	2	296	0,13	-0,01	0,41	
248	3,25	-2,48	0,20		297	0,15	-0,20	2,09	
249	0,52	0,14	-0,48	2	298	0,21	0,00	0,23	2
251	0,09	-0,07	-0,73	S	299	0,16	0,02	-0,14	2
252	0,15	0,06	-0,56	1	301	0,26	-0,03	-0,06	A
253	0,11	0,12	0,39	2	302	0,29	-0,07	1,59	
254	0,13	0,05	-0,16	2	303	0,27	0,05	-0,62	1
255	-0,45	0,19	-0,69		304	0,35	0,15	0,07	2
256	0,15	0,17	-0,23	2	305	0,29	0,03	0,06	A
257	0,02	0,21	-2,91	1	306	0,33	0,15	-1,50	1
258	0,12	-0,13	0,88		307	0,30	-0,33	2,57	
259	0,02	-0,16	0,34		308	0,29	0,06	-0,26	2
263	-0,10	0,13	0,05	2	309	0,18	-0,05	1,04	
264	0,19	0,56	-4,59	1	311	0,13	-0,03	0,32	
265	0,93	-0,28	-1,58	S	312	0,01	-0,01	0,92	
266	0,20	0,07	-0,28	2	313	0,36	-0,03	0,11	
267	0,22	0,07	-2,16	1	314	0,19	0,00	-0,90	1
269	0,23	-0,14	0,49		315	0,10	0,05	-0,18	2
271	0,15	0,09	-0,63	1	319	0,16	0,11	-0,19	2
272	0,36	-0,36	-0,02	S	321	0,15	0,00	-0,75	1
273	-0,25	0,27	-2,30		322	-0,02	0,29	-1,23	
274	0,04	0,08	-1,41	1	323	0,22	-0,05	1,28	
275	0,38	-0,25	1,95		324	0,26	-0,74	3,36	
279	0,18	0,11	-1,53	1	325	0,29	0,00	-0,23	S
281	0,11	0,10	-1,91	1	326	0,06	-0,06	1,64	
282	0,19	0,08	-0,38	2	327	0,23	0,04	0,46	2
283	0,20	0,02	-0,01	A	331	0,47	-1,11	1,83	
284	0,14	0,00	-0,17	2	341	0,09	-0,24	-0,44	S
285	0,08	0,01	-0,10	2	342	-0,02	-0,34	2,63	
286	0,19	-0,14	0,95		343	0,10	-0,13	0,40	
287	0,13	-0,10	1,31		344	0,14	0,01	-0,60	1
288	0,04	0,06	-0,66	1	345	0,21	0,08	-1,99	1
289	0,23	-0,06	-0,17	S	346	0,06	0,14	-1,58	1
291	0,11	-0,22	1,94		348	0,26	0,02	0,07	A
292	-0,01	0,17	-1,75		349	0,40	-0,11	0,09	
293	0,10	0,04	0,31	2					

5.4 Descriptive statistics: Table A3

1. 30 Peripherial Prefectures with Average of Core Prefectures

Variables	Observations	Mean	Std. Dev.	Min	Max
meangap	2644	0,217344	0,352457	-1,343929	2,374911
stdgap	2644	0,094221	0,395814	-2,592987	1,59148
skewgap	2644	-1,327319	1,353942	-5,14438	4,154283
std periphery	2644	0,746842	0,407578	0,053094	4,169486
skew periphery	2644	-0,600088	1,164951	-5,340727	2,57987
Dist	2644	5,451519	0,784737	3,427515	6,725274
Emp	2644	2,442071	0,502103	1,386294	4,942278
lkl	2644	3,604542	0,78518	-0,993252	8,060988

2. Sector Regressions

Variables	Observations	Mean	Std. Dev.	Min	Max
std	150	0,858653	0,27957	0,249146	2,730309
mean	150	6,162526	0,357962	5,057228	7,780544
std periphery	150	0,903392	0,396367	0,181375	3,859641
lvaemp	150	6,085545	0,391167	4,631279	7,819726
KL	150	6,085545	0,391167	4,631279	7,819726
Lower25	150	4,393508	1,92145	0	6,414497
emp	150	2,369816	0,476651	1,722575	5,365081

References

Axtell, R. L. (2001): “Zipf Distribution of U.S. Firm Sizes,”Science, 293.

Baldwin R., Forslid R., Martin P., Ottaviano G. and F. Robert-Nicoud, (2003). Economic geography and public policy, Princeton University Press.

Baldwin R. and T.Okubo, (2006). Heterogenous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting, Journal of Econiomic Geography 6(3), pp.323-346.

Bernard A., B. Jensen, S. Redding, and P. Schott, (2007). Firms in International Trade, The Journal of Economic Perspectives 21(3), pp. 105-130.

Cabral, L., and Mata, J. (2003) “On the evolution of the firm size distribution: facts and theory,” American Economic Review 93, 1075-1090.

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

Combes P-P and H.Overman, (2003). The spatial distribution of economic activities in the European Union, Handbook of Urban and Regional Economics, vol. 4, V. Henderson and J.-F. Thisse (eds.), Elsevier-North Holland, Amsterdam.

Dixit, A. and J. Stiglitz, (1977). "Monopolistic competition and optimum product diversity". American Economic Review 67: 297-308.

Del Gatto M. Ottaviano G.I.P. and M. Paganini, (2008). "Openness to Trade and Industry Cost Dispersion: Evidence from a Panel of Italian Firms".

Fukao, K. and X. Yue, (2000). "Economic convergence and production input in post-war Japan - An application of the Solow economic growth model", Keizai Kenkyu 52 (2).

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): “Export versus FDI with Heterogeneous Firms,”American Economic Review, 94(1), 300–316.

Luttmer, E. G. J. (2007): “Selection, Growth, and the Size Distribution of Firms,”Quarterly Journal of Economics, 122(3), 1103–1144.

Melitz, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, Econometrica, 71(6), 1695–1725.

Melitz, M. J., and G. I. P. Ottaviano (2005): “Market Size, Trade and Productivity,”NBER Working Paper No. 11393.

Martin, P. and C.A. Rogers, (1995). "Industrial location and public infrastructure", Journal of International Economics 39, 335-351.

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.

Okubo, T, Picard, P.M., and Thisse, J-F (2010) “The spatial selection of heterogeneous firms,” Journal of International Economics, forthcoming.

Okubo, T and E.Tomiura (2010a). "Industrial Relocation Policy and Heterogeneous Plants Sorted by Productivity: Evidence from Japan," Discussion papers 10016, Research Institute of Economy, Trade and Industry (RIETI).

Okubo, T and E.Tomiura, (2010b). "Productivity distribution, firm heterogeneity, and agglomeration: Evidence from firm-level data," Discussion papers 10017, Research Institute of Economy, Trade and Industry (RIETI).

Ottaviano G.I.P., T. Tabuchi and J-F Thisse (2002). "Agglomeration and Trade Revisited", International Economic Review 43, 409-36.

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. (2004): "Market Structure and Productivity: A Concrete Example,"Journal of Political Economy, 112(6).