WHICH FIRMS ARE LEFT IN THE PERIPHERY? SPATIAL SORTING OF HETEROGENEOUS FIRMS WITH SCALE ECONOMIES IN TRANSPORTATION*

Rikard Forslid
Department of Economics, Stockholm University, Stockholm 10691, Sweden.
E-mail: Rikard.Forslid@ne.su.se

Toshihiro Okubo
Faculty of Economics, Keio University, 2-15-45 Mita Minato-ku, Tokyo 108-8345, Japan.
E-mail: okubo@econ.keio.ac.jp

ABSTRACT. This paper introduces scale economies in transportation in a trade and geography model with heterogeneous firms. This relatively small change to the standard model produces a new pattern of spatial sorting among firms. In contrast to the existing literature, our model produces the result that firms of intermediate productivity relocate to the large core region, whereas high- and low-productivity firms remain in the periphery. Trade liberalization leads to a gradual relocation to the core with the most productive firms remaining in the periphery.

1. INTRODUCTION

This paper introduces scale economies in transportation in a trade and geography model with heterogeneous firms. It is shown how scale economies in transportation can dramatically change the location pattern of firms. Baldwin and Okubo (2006) show how the most productive firms tend to agglomerate in the large central region, when firm heterogeneity is added to a trade and geography framework.[1] Thus, their model produces a spatial sorting pattern with the most productive firms in the center and the least productive firms in the periphery. The sorting pattern is amplified by trade liberalization. Lower trade costs induce more firms to relocate to the core (central) region, and it is always the most productive firm in the periphery that has the strongest incentive to relocate towards the core. The result by Baldwin and Okubo (2006) underscores the concerns that small regions or countries could end up as losers when economic integration leads to an agglomeration of firms to large central regions. Not only does the periphery lose industry in general, it is the most productive firms that leave.

The reason for our result is that transportation costs become relatively less important for the large and most productive firms in our model. The reason for this is that they obtain low freight rates because of their large shipments. Furthermore, trade liberalization accentuates the spatial sorting pattern. Our results imply that different degrees of

*We are grateful for financial support from the Grant-in-Aid for Scientific Research (JSPS) and Research Institute of Economy, Trade and Industry (RIETI), the Jan Wallander and Tom Hedelius' Research Foundation, and The Swedish Research Council.

Received: March 2013; revised: December 2013; accepted: December 2013.

[1]Similar results are found by Okubo, Picard, and Thisse (2010) using the linear model of Ottaviano, Tabuchi, and Thisse (2002) and by Saito, Gopinath, and Wu (2011) and Forslid and Okubo (2012) using a three-country model. The vertical linkage model also produces similar results (Okubo, 2009).

 DOI: 10.1111/jors.12115

scale economies among industrial sectors lead to different spatial sorting patterns among sectors.

Our analysis shows that scale economies in transportation produces a countervailing force to the tendency for the most productive firms to locate in central locations. It is possible that this effect is dwarfed by other forces in reality. Nevertheless, using Japanese plant level data, Fukao et al. (2011) find that productive firms are likely to set up new factories in the periphery rather than in core regions. Likewise, using the same Japanese data, Okubo and Tomiura (2013) show that a high proportion of very highly productive exporting firms are found in peripheral areas alongside a large proportion of low-productivity nonexporting firms. These findings are consistent with the spatial sorting pattern in our paper.

Our results are driven by firm level differences in transport costs, and these remain substantial in spite of the globalization process during the last decades. In fact, the tramp freight rates have roughly remained constant since the 1950s, while the airfreight rates have fallen (Hummels, 2007). Anderson and van Wincoop (2004) find that the total trade costs in high income countries amount to an ad valorem tax equivalent to 170 percent. There are several strands of literature that analyze trade costs. One strand deals with the Alchian Allen conjecture or the "shipping the good apples out" theorem, stating that trade costs tend to be lower in ad valorem terms for high-quality goods when trade costs are specific (per unit), and that this leads to more trade in high-quality goods (Borcherding and Silberberg, 1978; Hummels and Skiba, 2004). Another strand analyses density economics in transportation, whereby intensely used trade routes have lower freight rates (e.g., Behrens et al., 2006).

This paper belongs to the trade and economic geography literature (see e.g. Fujita et al., 1999; Baldwin et al., 2003; Fujita and Thisse, 2002), which typically treats trade costs as a constant parameter. As pointed out by McCann (2001, 2005), this is at odds with wide-ranging empirical work on transport rate structures in regional economics which has almost universally found economies of distance and scale in transportation. There are a few trade and geography papers that do model transport costs in an explicit manner. Bosker and Garretsen (2010) show how the formulation of the trade cost function affects the empirical relevance of market access in determining wage levels. Behrens et al. (2006) introduce density economies in transportation in a trade and geography model. They show how this may give rise to multiple equilibria and a catastrophic agglomeration of industry. Behrens and Picard (2011) study the location choice of firms when the transportation costs are endogenous. They show how the back-haul problem increases the freight rates on exports from the central country, and that this dampens the tendencies for agglomeration. However, none of these papers analyze spatial sorting of firms, as they work with models with homogenous firms.

Scale Economies in Transportation

We will here relate firms' varying production volumes to differences in transportation costs. It is quite natural that freight rates per unit decline with the size of a shipment due to scale economies in transportation. Scale economies in transportation are well established in regional economics (see e.g., McCann, 2005), and Hummels and Skiba (2004) find that a 10 percent increase in product weight/value leads to a 4–6 percent increase in shipping costs using U.S. trade data.[2] However, almost all types of

[2]Another example is Mori and Nishikimi (2002) who find that the cost of transportation by ship declines by 0.31 percent when the shipping volume increases by 1 percent.

transportation such as trucks, airfreight, and ships have freight rates that are declining in the volume transported.

Note that there are scale economies in transportation at different levels. First, the handling of a large shipment is generally more efficient than the handling of several small shipments. Firm-level freight rates are therefore declining in the volume transported. Thus, there are scale economies in transportation at the exporter firm level. Second, at a more aggregate level, the freight rates for densely used trade routes tend to be lower, e.g., because larger and more efficient ships or airplanes can be used on such routes. This type of scale economies is often named density economics in transportation. In this paper, we will concentrate on the former type of scale economies in transportation. We use a set-up with heterogeneous firms and two regions, which implies that the more interesting variation is at the firm level.

2. MODEL

Here, we introduce scale economies in transportation in the Baldwin and Okubo (2006) model. This model is based on the Melitz (2003) heterogeneous firms trade model combined with the "footloose capital" new economic geography model by Martin and Rogers (1995). Our model involves firm heterogeneity in transport costs as well as in labor productivity. In essence, productive firms export more and their trade costs per unit are therefore lower than those of less productive firms due to scale economies in transportation.

Basics

There are two regions with an asymmetric population (market size). Region j is the larger core region and Region k is a smaller region. There are two types of factors of production, capital and labor, and each country has the same proportion of capital to labor. That is, regions are identical except for size and Region j has a share s_j of both factors of production. Capital can move between regions but capital owners do not. Labor can move freely between sectors but are immobile between regions. A homogeneous good is produced with a constant-returns technology only using labor. Differentiated manufactures are produced with increasing-returns technologies using both capital and labor.

All individuals have the utility function

$$U = C_M^{\mu} C_A^{1-\mu}, \quad \text{where} \quad C_M = \left[\int_{l \in \Psi} c_l^{(\sigma-1)/\sigma} dl \right]^{\sigma/(\sigma-1)}, \tag{1}$$

where $\mu \in (0, 1)$, $\sigma > 1$ are constants, and Ψ is the set of consumed variety. C_M is a consumption index of manufacturing goods and C_A is consumption of the homogeneous good. c_l is the amount consumed of variety l. Region subscripts are suppressed when possible for ease of notation.

Each consumer spends a share μ of his income on manufactures. The total demand for a domestically produced variety i is

$$x_i = \frac{p_i^{-\sigma}}{\int_{l \in \Psi} p_l^{1-\sigma} dl} \cdot \mu Y, \tag{2}$$

where p_l is the price of variety l and Y is income in the region.

Ownership of capital is assumed to be fully regionally 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. Therefore, the income of each region is constant and independent of the location of capital. Total expenditure equals total factor income. Firms' fixed factor of production is capital and the variable factor is labor. The return to capital therefore equals firms' operating profit in equilibrium. Thus, the total equilibrium expenditures can be written $E = wL + rK = wL + \mu E/\sigma$. Without loss of generality, we choose units so that $L \equiv 1$, which gives $E = \frac{1}{1-\mu/\sigma}$. The income of Region j is equal to its share of total expenditures given by

$$(3) \qquad Y_j = s_j E = s_j \frac{\sigma}{\sigma - \mu}.$$

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

Turning to the supply side, the homogeneous good sector is a constant returns and perfect competition sector. The unit factor requirement of the homogeneous good is one unit of labor. The good is freely traded and since it is also chosen as the numeraire, we have

$$(4) \qquad p_A = w = 1,$$

where w is the wage of workers in all regions.

In the production of differentiated goods, firms have a firm-specific unit labor input coefficient (a) and uses one unit of capital, as in the standard footloose capital model. There is a fixed world capital endowment, implying that the world mass of firms, N^W, is constant, whilst international capital mobility allows firms to move between regions. We normalize the world mass of differentiated firms to one, $N^W \equiv 1$. The total costs for firm i are specified as

$$(5) \qquad TC_i = \pi_i + a_i x_i,$$

where π is return to capital. The variable cost consists of labor. Importantly, firms are heterogeneous and their firm-specific marginal production costs a_i are distributed according to the cumulative distribution function $G(a)$.

Geographical distance is represented by trade costs. Shipping the manufactured good involves a frictional trade cost of the "iceberg" form: for one unit of good from firm i in Region j to arrive in Region k, $\tau_{ijk} > 1$ units must be shipped. The trade costs are symmetric between regions $\tau_{ijk} = \tau_i \; \forall \; j, k$.

Profit maximization by manufacturing firms leads to a constant mark-up over the marginal cost

$$(6) \qquad p_i = \frac{\sigma}{\sigma - 1} a_i,$$

and the export price is $p_i \tau_i$.

We assume that there are scale economies in transportation, implying that the freight rate falls with the quantity shipped. Firms are heterogeneous in labor productivity and productive firms produce more and export more. Demand for firm i's product is proportional to $a_i^{-\sigma}$ from (2) and (6). Thus, productive firms (firms with a low a) sell larger quantities in all markets and therefore, they also export larger quantities which, in turn, implies lower freight rates. Therefore, we use firm productivity as a proxy for the firm level freight rate in this section of the paper, whereas we turn to the simulation of a more general case below. We here assume the following relationship between transport costs (freight rates) and firm-level productivity a_i:

$$\tau_i = t a_i^{\gamma}, \tag{7}$$

where $t \geqslant 1$ and $\gamma \geqslant 0$ are parameters.[3] t represents the fundamental level of the iceberg trade cost, whereas γ is the degree of scale economies of transportation. A larger γ implies stronger scale economies. We assume that $\gamma > 0$, implying that $\frac{d\tau_i}{da_i} > 0$. Note that our specification generates the standard case of a constant trade cost ($\tau_i = t$) when $\gamma = 0$.

Firm heterogeneity in labor requirements, a_i, is probabalistically allocated among firms. To analytically solve the model, we follow Forslid and Okubo (2012) and assume the following cumulative density function of a:

$$G(a) = \frac{a^{\rho} - \underline{a}^{\rho}}{a_0^{\rho} - \underline{a}^{\rho}}, \tag{8}$$

where ρ is a shape parameter and a_0 is a scaling factor. We assume the distribution to be truncated at $\underline{a}$, which is exogenously given, meaning that $\underline{a} \leq a \leq a_0$. This implies that the productivity of firms is bounded. Without loss of generality, we can normalize $a_0 \equiv 1$. For convenience, the lower bound for the marginal cost is specified as $\underline{a} \equiv t^{-\frac{1}{\gamma}}$, which implies that $\tau_i = 1$ when $a = \underline{a}$.[4]

Short-Run Equilibrium

In the short-run equilibrium, the allocation of capital in each region is taken to be fixed. s_j denotes the share of capital and the number (mass) of firms in Region j, since one unit of capital corresponds to one firm.

The return to capital of a firm in Region j is the firm's operating profit,

$$\pi_j(a_i) = \frac{a_i^{1-\sigma}}{(\sigma - \mu)} \mu \left(\frac{s_j}{\overline{\Delta}_j} + \phi_i \frac{1 - s_j}{\overline{\Delta}_k} \right), \tag{9}$$

where $\phi_i \equiv \tau_i^{1-\sigma}$, the right-hand side follows from the demand functions in (2), and

$$\overline{\Delta_j} \equiv s_j \int_{\underline{a}}^{1} a_i^{1-\sigma} dG(a) + (1 - s_j) \int_{\underline{a}}^{1} \phi_i a_i^{1-\sigma} dG(a). \tag{10}$$

The object ϕ_i, ranging between 0 and 1, stands for firm i "free-ness" of trade between Region j and k (0 is autarky and 1 is zero trade costs). It is assumed that the labor stock is sufficiently large so that the homogenous sector, which pins down the wage, is active in all regions.

Furthermore, using (7) and defining $\varphi \equiv t^{1-\sigma} \in [0, 1]$ gives $\phi_i = \tau_i^{1-\sigma} = \varphi a_i^{\gamma(1-\sigma)}$, where $a_i^{\gamma(1-\sigma)}$ is a firm-specific component and φ is a measure of the overall degree of trade freeness.

---

[3]Our specification, which makes the model analytically tractable, is a simplification of the more general formulation $\tau = \tau(t, x(a))$, with $\tau_x{}' < 0$. Below, we numerically simulate the case where trade costs are directly related to the export volume. The simulations produce qualitatively the same results as the analytically tractable case.

[4]The assumption that the most productive firm has zero trade costs is convenient when solving the model, and it rules out negative trade costs.

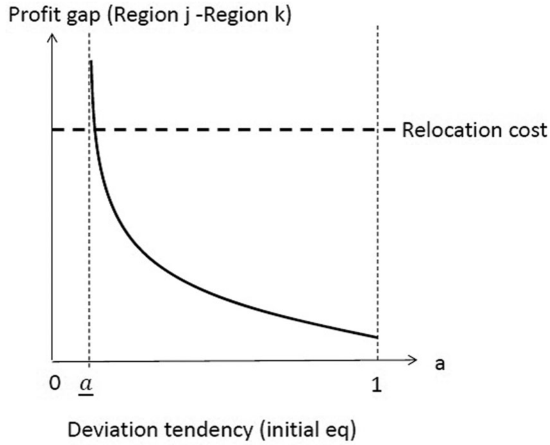

FIGURE 1: No Scale Economies in Transportation.

Consider now what would happen if firms were allowed to move between regions. Firms choose location in order to obtain the highest capital returns. Using (9), a firm will move from the larger core Region j to Region k when

$$\pi_j(a_i) - \pi_k(a_i) - \chi = \frac{a_i^{1-\sigma}}{(\sigma - \mu)}(1 - \phi_i)\mu\left(\frac{s_j}{\overline{\Delta}_j} - \frac{1 - s_j}{\overline{\Delta}_k}\right) - \chi > 0, \tag{11}$$

where χ is a per-unit of capital fixed relocation cost and trade costs are assumed to be symmetric, $\tau_{jk} = \tau_{kj}$.

Relocation Tendencies

Before moving to the full long-run equilibrium of the model, we consider the relocation incentives faced by firms starting out from the initial equilibrium. Using (7), the profit differential is given by

$$\pi_j(a_i) - \pi_k(a_i) - \chi = \frac{\mu}{(\sigma - \mu)}(1 - \varphi a_i^{\gamma(1-\sigma)})a_i^{1-\sigma}\left(\frac{s_j}{\overline{\Delta}_j} - \frac{1 - s_j}{\overline{\Delta}_k}\right) - \chi. \tag{12}$$

Differentiating the profit gap function in terms of a_i gives:

$$\frac{d(\pi(a_i) - \pi(a_i))}{da_i} = -(1 - \varphi(1 + \gamma)a_i^{\gamma(1-\sigma)})a_i^{-\sigma}\frac{\mu(\sigma - 1)}{(\sigma - \mu)}\left(\frac{s_j}{\overline{\Delta}_j} - \frac{1 - s_j}{\overline{\Delta}_k}\right). \tag{13}$$

From this expression, it appears that the incentives of firms depend on the level of scale economies in transportation, γ. With no scale economies, $\gamma = 0$, we have returned to the Baldwin and Okubo (2006) case as depicted in Figure 1. Only the most productive firms (the low a firms) will have incentives to relocate in this case. The most productive firms are better equipped to handle the strong competition in the large market, and only these firms will expand the sales and increase the profits sufficiently when moving to overcome the relocation cost. $\gamma = 0$ therefore tends to lead to spatial sorting with the most productive firms sorting to the larger market (Region j).

The relocation tendencies are very different when there are scale economies in transportation, $\gamma > 0$. High scale economies and not too low trade costs ($\gamma > \frac{1}{\varphi} - 1$) will lead to the opposite case of a positively sloping profit differential curve as seen from (13).

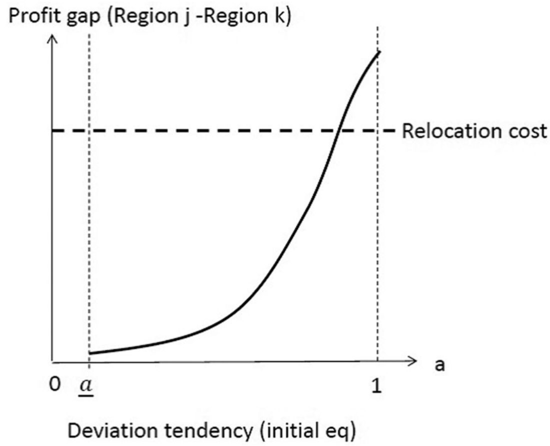

FIGURE 2: High Scale Economies in Transportation.

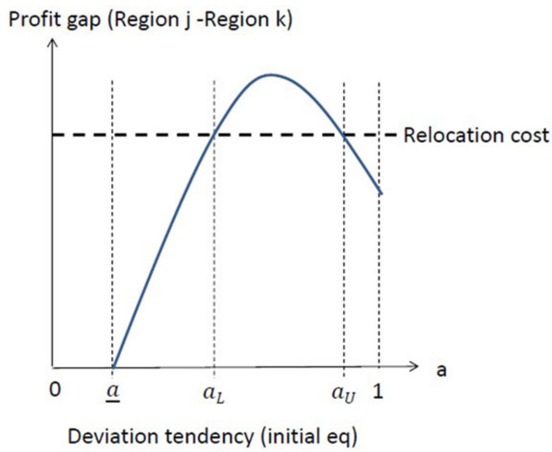

FIGURE 3: The Generic Case of Intermediate Scale Economies in Transportation ($\gamma < \frac{1}{\varphi} - 1$).

Figure 2 illustrates this case. Here, it is the least productive firms that have the strongest incentives to relocate to the larger Region j. High-productivity firms enjoy low transportation costs, and location therefore becomes relatively unimportant for them. Low-productivity firms, in contrast, face high transportation costs and therefore, they have large gains from moving. Thus, this case tends to lead to spatial sorting with high-productivity firms in the periphery.

Finally, the generic case, when scale economies in transportation are not too large, is a hump-shaped profit differential curve as shown in Figure 3. The exact conditions for this case are stated in the following proposition:

PROPOSITION 1. *The firms of intermediate productivity have the strongest incentives to start relocating to the large region if the conditions* $\gamma < \frac{1}{\varphi} - 1$, *and* $\gamma > 0$ *hold.*

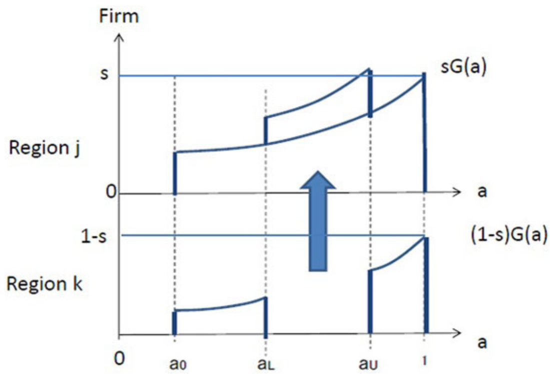

FIGURE 4: The Relocation Pattern with Two Cut-Offs.

First, when $a = \underline{a} \equiv t^{-\frac{1}{\gamma}}$, $\frac{d(\pi(a_i)-\pi(a_i))}{da} > 0$ always holds from (13). On the other hand, from (13), when $a = 1$, $\gamma > 0$, and $\gamma < \frac{1}{\varphi} - 1$, $\frac{d(\pi(a_i)-\pi(a_i))}{da} < 0$. This implies that the profit gap curve, $\pi(a_i) - \pi(a_i)$, must be hump-shaped.

We will assume that the sufficient conditions in Proposition 1 hold in the following. Intuitively, the least productive firms have difficulties surviving in the large region, while the most productive firms have such low transportation costs that they do not find it profitable to pay the relocation cost in order to move to the core. Therefore, it is the firms with intermediate productivity that have the strongest incentives to move from the periphery to the core. This relocation tendency stands in sharp contrast to that in Baldwin and Okubo (2006) and other studies in this vein, where the most productive firms always have the largest incentive for relocation and are the first movers to the large market.

Having analyzed the relocation tendencies starting from a short-run equilibrium where capital is fixed, we now turn to the long-run equilibrium after firms and capital have relocated.

The Long-Run Equilibrium

Capital is fully mobile between regions in the long run. To simplify the analysis, we assume that there is a firm (capital) relocation cost à la Baldwin and Okubo (2006) that is related to the migration pressure. More specifically, we assume the cost of moving to be given by

$$\chi = \kappa\left(\frac{dF(a_U)}{dt} - \frac{dF(a_L)}{dt}\right), \tag{14}$$

where κ is a positive constant, and a_U and a_L are the two cut-off productivities for the hump-shaped profit differential curve (see Figure 3). The relocation cost is high when many firms move at the same time, but gradually declines to zero as the migration pressure falls as we approach the equilibrium. The firm with the most to gain from relocating has the highest willingness to pay the cost for relocation. As a consequence, firms relocate in order of productivity. The relocation process is illustrated by Figure 4.

Firms move as long as the profit gap is larger than the relocation cost, and the long-run equilibrium is therefore defined by:

$$\pi_j(a_i) - \pi_k(a_i) = \frac{\mu}{(\sigma-\mu)}(1 - \varphi a_i^{\gamma(1-\sigma)})a_i^{1-\sigma}\left(\frac{s_j}{\Delta_j} - \frac{1-s_j}{\Delta_k}\right) = 0. \tag{15}$$

The first thing to note, from (15), is that the profit differential is zero for $a = \underline{a} \equiv t^{-\frac{1}{\gamma}}$. This implies that $a_L = \underline{a} \equiv t^{-\frac{1}{\gamma}}$ in a long-run equilibrium. Next, the long-run equilibrium value of a_U is defined by

$$\text{(16)} \qquad \pi_j(a_U) - \pi_k(a_U) = \frac{\mu}{(\sigma - \mu)}(1 - \varphi\, a_U^{\gamma(1-\sigma)})a_U^{1-\sigma}\left(\frac{s_j}{\Delta_j} - \frac{1-s_j}{\Delta_k}\right) = 0,$$

where

$$\text{(17)}\ \Delta_j \equiv \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left(s_j\int_{\underline{a}}^{1} a^{1-\sigma}dF(a) + (1-s_j)\int_{\underline{a}}^{1}\phi a^{1-\sigma}dF(a) + (1-s_j)\int_{a_L}^{a_U}(1-\phi)a^{1-\sigma}dF(a)\right),$$

$$\text{(18)}\ \Delta_k \equiv \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\left((1-s_j)\int_{\underline{a}}^{1} a^{1-\sigma}dF(a) + s_j\int_{\underline{a}}^{1}\phi a^{1-\sigma}dF(a) + (1-s_j)\int_{a_L}^{a_U}(\phi-1)a^{1-\sigma}dF(a)\right).$$

Solving (17) and (18) gives

$$\text{(19)} \quad \Delta_j = \lambda\{s_j(1-\underline{a}^{\beta}) + (1-s_j)(a_U^{\beta} - a_L^{\beta})\} + \theta\varphi\{(1-s_j)(a_L^{\alpha} - \underline{a}^{\alpha}) + (1-s_j)(1-a_U^{\alpha})\},$$

$$\text{(20)} \quad \Delta_k = \theta\varphi\{s_j(1-\underline{a}^{\alpha}) + (1-s_j)(a_U^{\alpha} - a_L^{\alpha})\} + \lambda\{(1-s_j)(a_L^{\beta} - \underline{a}^{\beta}) + (1-s_j)(1-a_U^{\beta})\},$$

where $\alpha \equiv (\gamma+1)(1-\sigma) + \rho$, $\beta \equiv 1 - \sigma + \rho$, $\lambda \equiv \frac{\rho}{\beta}\frac{1}{a_0^{\rho} - \underline{a}^{\rho}}$, and $\theta \equiv \frac{\rho}{\alpha}\frac{1}{a_0^{\rho} - \underline{a}^{\rho}}$. Note that $\alpha < \beta$ (since $\sigma > 1$) and thus, $\lambda < \theta$. Moreover, we note that $\beta\lambda = \alpha\theta$.

The relevant solution for a_U is determined by $\frac{s_j}{\Delta_j} - \frac{1-s_j}{\Delta_k} = 0$. Manipulation gives the following condition that determines the upper cut-off, a_U:

$$\text{(21)} \qquad \frac{\alpha}{\beta}a_U^{\beta} - \varphi a_U^{\alpha} = \frac{\alpha}{\beta}t^{-\frac{\beta}{\gamma}} - \varphi\frac{s_j}{1-s_j}t^{-\frac{\alpha}{\gamma}} - \varphi(1 - \frac{s_j}{1-s_j}) = 0.$$

This condition cannot be solved analytically in general but we can use it to derive comparative statistics on a_U, as seen below. We next turn to the effect of trade liberalization.

Trade Liberalization

PROPOSITION 2. *Trade liberalization always promotes firm relocation from small to large regions. The most productive firms are the last movers.*

Although we cannot derive a closed form solution for the cut-off, we can investigate the impact of trade liberalization, i.e., the impact of an increase in φ on the cut-off, by implicit differentiation. Using that $\varphi = t^{1-\sigma}$ in (21) gives

$$\text{(22)} \qquad F \equiv \frac{\alpha}{\beta}a_U^{\beta} - \varphi a_U^{\alpha} = \frac{\alpha}{\beta}\varphi^{\frac{\beta}{\eta}} - \varphi^{1-\frac{\alpha}{\eta}}\frac{s_j}{1-s_j} - \varphi\left(1 - \frac{s_j}{1-s_j}\right) = 0,$$

where $\eta \equiv \gamma(\sigma - 1)$.

First

$$\text{(23)} \qquad \frac{dF}{da_U} = \alpha a_U^{\beta-1} - \varphi\alpha a_U^{\alpha-1} > 0$$

since $\beta > \alpha$ and $\varphi < 1$.

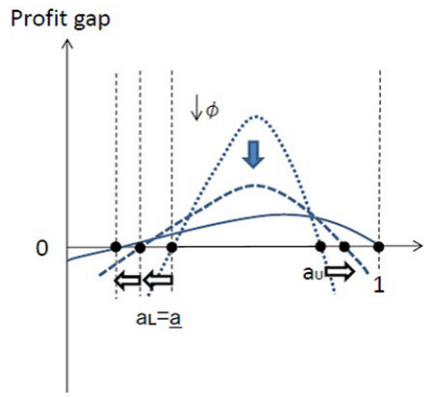

FIGURE 5: The Effect of Trade Liberalization.

Next

$$\frac{dF}{d\varphi} = -a_U^{\alpha} - \frac{\alpha}{\eta}\varphi^{\frac{\beta}{\eta}-1} + (1-\frac{\alpha}{\eta})\varphi^{-\frac{\alpha}{\eta}}\frac{s_j}{1-s_j} + \frac{1-2s_j}{1-s_j}. \tag{24}$$

All terms on the right-hand side of this expression are negative assuming that $\frac{\alpha}{\eta} > 1$. Using that $\alpha \equiv (1+\gamma)(1-\sigma)+\rho$, this condition may be rewritten as $\gamma < \frac{\sigma-1}{\rho-\sigma+1}$, which implies that

$$\frac{dF}{d\varphi} < 0. \tag{25}$$

Finally, using the implicit function theorem

$$\frac{da_U}{d\varphi} = -\frac{\frac{dF}{d\varphi}}{\frac{dF}{da_U}} > 0. \tag{26}$$

Thus, we have shown that, as the trade costs fall, more firms move from the smaller Region k to the larger Region j, leaving an increasingly smaller set of very high-productivity firms in the periphery. The effect of trade liberalization is illustrated by Figure 5.

Welfare Impact of Trade Liberalization

PROPOSITION 3. *Trade liberalization always improves welfare in the large region, while the welfare effect is ambiguous for the small region.*

Using the Cobb–Douglas utility function, social welfare in our model in the large Region j and the smaller Region k is given by $U_j = \frac{s_j}{\Delta_j^{\mu/(1-\sigma)}}$ and $U_k = \frac{1-s_j}{\Delta_k^{\mu/(1-\sigma)}}$. Our central interest is Δ_j and Δ_k. First, as shown in the Appendix.

$$\frac{d\Delta_j}{d\varphi} = \frac{d\Delta_j}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_j}{da_L}\frac{da_L}{d\varphi} + \frac{d\Delta_j}{d\varphi} > 0. \tag{27}$$

Thus, firm relocation improves social welfare in Region j. Then, the same investigation is conducted for Region k. The welfare effects are ambiguous in this case (see the Appendix).

$$\frac{d\Delta_k}{d\varphi} = \frac{d\Delta_k}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_k}{da_L}\frac{da_L}{d\varphi} + \frac{d\Delta_k}{d\varphi} \lesseqgtr 0.$$

The indirect effect is always negative, i.e., $\frac{d\Delta_k}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_k}{da_L}\frac{da_L}{d\varphi} < 0$, whereas the direct effect is always positive $\frac{d\Delta_k}{d\varphi} > 0$. This indicates that if trade liberalization does not relocate firms to any large extent, i.e., because of large relocation costs, welfare is likely to rise. For low relocation costs, the opposite is likely to occur.

3. FIRM LEVEL TRADE COSTS AS A FUNCTION OF THE EXPORT VOLUME

In order to handle the model analytically, we have so far used firm productivity as a proxy for firm shipments when modeling firm-level trade costs. We now turn to the case when the firm-level trade cost decreases in the quantity it exports. The firm i transport cost from Region j to Region k is given by

$$\tau_{ijk} = t x_{ijk}^{-\gamma}, \tag{28}$$

where x_{ijk} is the quantity transported and γ is the elasticity of the trade cost w.r.t. the export quantity. The first thing to note is that this specification implies that trade costs are asymmetric because of the asymmetric market sizes. Using (2) and (28) gives

$$\phi_{ijk} = \tau_{ijk}^{1-\sigma} = a_i^{-\frac{\sigma\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} t^{\frac{1-\sigma}{1-\gamma(\sigma-1)}} \left(\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \frac{\mu\frac{\sigma}{\sigma-\mu}(1-s_j)}{\Delta_k} \right)^{\frac{\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} = a_i^{-\frac{\sigma\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} \Theta B_k^{\frac{\gamma(\sigma-1)}{1-\gamma(\sigma-1)}}, \tag{29}$$

$$\phi_{ikj} = \tau_{ikj}^{1-\sigma} = a_i^{-\frac{\sigma\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} t^{\frac{1-\sigma}{1-\gamma(\sigma-1)}} \left(\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \frac{\mu\frac{\sigma}{\sigma-\mu}s_j}{\Delta_j} \right)^{\frac{\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} = a_i^{-\frac{\sigma\gamma(\sigma-1)}{1-\gamma(\sigma-1)}} \Theta B_j^{\frac{\gamma(\sigma-1)}{1-\gamma(\sigma-1)}}, \tag{30}$$

where $\Theta \equiv t^{\frac{1-\sigma}{1-\gamma(\sigma-1)}}((\frac{\sigma}{\sigma-1})^{-\sigma}\mu\frac{\sigma}{\sigma-\mu})^{\frac{\gamma(\sigma-1)}{1-\gamma(\sigma-1)}}$, $B_k \equiv \frac{1-s_j}{\Delta_k}$, and $B_j \equiv \frac{s_j}{\Delta_j}$. We wish to rule out negative trade costs and therefore assume that $\underline{a} \geq \Theta^{\frac{1-\gamma(\sigma-1)}{\sigma\gamma(\sigma-1)}} B_l^{\frac{1}{\sigma}}, l \in \{j,k\}$, which limits the productivity of the most productive firm.

Before turning to simulations, we analyze the deviation tendencies starting out from a situation where firms have not yet moved. The profit gap between the two markets is given by

$$\pi_j(a_i) - \pi_k(a_i) = \frac{a_i^{1-\sigma}}{(\sigma-\mu)}\mu\left(\frac{s_j}{\overline{\Delta}_j} + \phi_{ijk}\frac{1-s_j}{\overline{\Delta}_k} - \phi_{ikj}\frac{s_j}{\overline{\Delta}_j} - \frac{1-s_j}{\overline{\Delta}_k}\right), \tag{31}$$

where $\overline{\Delta}$ is given by (10). Substituting in (29) and (30) gives

$$\pi_j(a_i) - \pi_k(a_i) = \frac{a_i^{1-\sigma}}{(\sigma-\mu)}\mu\left(\overline{B_j} - \overline{B_k}\right) - \mu\Theta a_i^{-\frac{(\sigma-1)(1+\gamma)}{1-\gamma(\sigma-1)}}\left(\overline{B_j}^{\frac{1}{1-\gamma(\sigma-1)}} - \overline{B_k}^{\frac{1}{1-\gamma(\sigma-1)}}\right). \tag{32}$$

Differentiation w.r.t. a_i gives

$$\begin{aligned} \frac{d(\pi_j(a_i) - \pi_k(a_i))}{da_i} &= (1-\sigma)\frac{a_i^{-\sigma}}{(\sigma-\mu)}\mu\left(\overline{B_j} - \overline{B_k}\right) + \mu\Theta\frac{(\sigma-1)(1+\gamma)}{1-\gamma(\sigma-1)}a_i^{-\frac{\sigma}{1-\gamma(\sigma-1)}} \\ &\quad \left(\overline{B_j}^{\frac{1}{1-\gamma(\sigma-1)}} - \overline{B_k}^{\frac{1}{1-\gamma(\sigma-1)}}\right). \end{aligned} \tag{33}$$

First, note that $\overline{B_j} > \overline{B_k}$ before relocation (in the short run) since Region j is larger. The incentives to relocate are highest for high-productivity firms ($\frac{d(\pi_j(a_i)-\pi_k(a_i))}{da_i} < 0$) for low γ ($\gamma < \frac{1}{\sigma-1}$) as seen from (33). For very high γ, in contrast, the incentives to relocate are highest for low-productivity firms ($\frac{d(\pi_j(a_i)-\pi_k(a_i))}{da_i} > 0$). Finally, it is possible to find intermediate γ's for which the profit gap is hump-shaped in a. Thus, the deviation

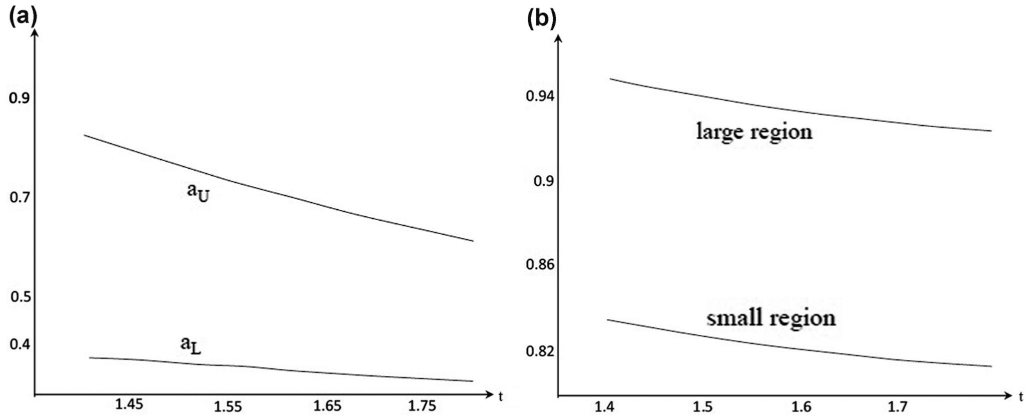

FIGURE 6: (a) The Effect of Trade Liberalization on the Cut-Offs. (b) Trade Liberalization and Welfare.

tendencies are very similar to that of the simpler specification of transport costs in the analytical section.

The analysis of the long-run equilibrium must be done by numerical simulation when the transport cost is specified according to (28). In particular, the price indices become more complex. Substituting (29) and (30) into expressions (17) and (18) gives, after some manipulation,

$$(34)\quad \Delta_j \equiv \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\frac{\rho}{1-\underline{a}^{\rho}}\left(\frac{s_j}{\rho-\sigma+1}\left[1-\underline{a}^{\rho-\sigma+1}\right]+\frac{(1-s_j)^{\frac{1}{1-\eta}}\Theta}{\rho-\frac{\sigma-1+\eta}{1-\eta}}\left(\frac{1}{\Delta_k}\right)^{\frac{\eta}{1-\eta}}\right.$$
$$\left.\left[1-\underline{a}^{\rho-\frac{\sigma-1+\eta}{1-\eta}}-a_U^{\rho-\frac{\sigma-1+\eta}{1-\eta}}+a_L^{\rho-\frac{\sigma-1+\eta}{1-\eta}}\right]+\frac{(1-s_j)}{\rho-\sigma+1}\left[a_U^{\rho-\sigma+1}-a_L^{\rho-\sigma+1}\right]\right),$$

and

$$(35)\quad \Delta_k = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma}\frac{\rho}{1-\underline{a}^{\rho}}\left(\frac{1-s_j}{\rho-\sigma+1}\left[1-\underline{a}^{\rho-\sigma+1}-a_U^{\rho-\sigma+1}+a_L^{\rho-\sigma+1}\right]\right.$$
$$\left.+\frac{s_j^{\frac{1}{1-\eta}}\Theta}{\rho-\frac{\sigma-1+\eta}{1-\eta}}\left(\frac{1}{\Delta_j}\right)^{\frac{\eta}{1-\eta}}\left[1-\underline{a}^{\rho-\frac{\sigma-1+\eta}{1-\eta}}\right]+\frac{(1-s_j)\,\Theta}{\rho-\frac{\sigma-1+\eta}{1-\eta}}\left(\frac{s_j}{\Delta_j}\right)^{\frac{\eta}{1-\eta}}\left[a_U^{\rho-\frac{\sigma-1+\eta}{1-\eta}}-a_L^{\rho-\frac{\sigma-1+\eta}{1-\eta}}\right]\right).$$

We simulate the model using as the equilibrium condition that the profit of firms are identical in the two regions at the cut-offs. Figures 6(a) and (b) show the result of trade liberalization in a typical simulation.Trade liberalization leads to an upward movement of the higher cut-off, a_U, in accordance with Proposition 2. Trade liberalization also means that a wider range of firms move towards the larger region (the gap between a_L and a_U widens). Also, trade liberalization leads to higher welfare in the large region in accordance with Proposition 3. Finally, welfare also increases in the small region in the simulations when trade is liberalized, whereas the effect is ambiguous in the analytical part.

4. CONCLUSION

This paper introduces scale economies in transportation in a trade and geography model with heterogeneous firms. This relatively small change to the standard set-up has large effects on the location pattern of firms. The seminal model by Baldwin and Okubo (2006), as well as other models in this vein, predicts that the most productive firms will sort themselves to the core region. Our paper overturns this result. Firms of intermediate productivity typically sort themselves to the core, whereas high- and low-productivity firms remain in the periphery. The result of Baldwin and Okubo (2006) only applies when the scale economies in transportation are very low.

We show how trade liberalization accentuates the sorting, gradually leaving only the most productive firms in the periphery. We also show that, whereas the central region gains from trade liberalization, the periphery may lose.

An implication of our results is that the spatial sorting of firms in a sector will depend on the degree of scale economies in transportation in that sector. This has several policy implications. Suppose, e.g., that the government prompts an industry to change from road transport to trains for environmental reasons. According to our model, this will affect the spatial sorting of the industry, if road transport and trains have different degrees of scale economies. However, the transportation pattern changes as firms move, and there is no guarantee that the new sorting pattern of firms implies less transportation. For example, if scale economies in transportation are more pronounced when it comes to rail transports, the large and most productive firms have less of an incentive to move to the core after switching to rail transport. This, in turn, implies a geographically more dispersed production, since fewer firms move to the core. Finally, a more geographically dispersed production could lead to more transportation, which is detrimental to the environment. Thus, our paper points out a new general equilibrium effect that may be relevant for several policies directed towards the transportation sector.

The trade and geography literature with homogenous firms shows how the agglomeration pattern of a sector depends on the level of transport costs. The introduction of firm heterogeneity implies that it is possible to analyze which firms that agglomerate; that is, it is possible to analyze the spatial sorting of firms. Our paper adds to this literature by showing that scale economies of transportation can have an important effect on the spatial sorting pattern of firms.

APPENDIX: WELFARE EFFECTS OF TRADE LIBERALIZATION

First we show that

$$\frac{d\Delta_j}{d\varphi} = \frac{d\Delta_j}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_j}{da_L}\frac{da_L}{d\varphi} + \frac{d\Delta_j}{d\varphi} > 0. \tag{A1}$$

From (17) to (18)

$$\frac{d\Delta_j}{da_U} = (1-s)\lambda\beta a_U^{\beta-1} - \varphi(1-s)\alpha\theta a_U^{\alpha-1} > 0, \tag{A2}$$

$$\frac{d\Delta_j}{da_L} = -(1-s)\lambda\beta a_L^{\beta-1} + \varphi(1-s)\alpha\theta a_L^{\alpha-1} < 0, \tag{A3}$$

$$\frac{d\Delta_j}{d\varphi} = \theta\{(1-s)(a_L^{\alpha} - \underline{a}^{\alpha}) + (1-s)(1-a_U^{\alpha})\} > 0. \tag{A4}$$

Since $\beta > \alpha$, $\beta\lambda = \alpha\theta$ and $\varphi < 1$. From (26) and since $a_L = t^{1-\sigma}$, in the long-run equilibrium we have

$$\frac{da_U}{d\varphi} > 0, \tag{A5}$$

$$\frac{da_L}{d\varphi} < 0. \tag{A6}$$

Together, these derivatives imply that $\frac{d\Delta_j}{d\varphi} > 0$.

Next, we illustrate that

$$\frac{d\Delta_k}{d\varphi} = \frac{d\Delta_k}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_k}{da_L}\frac{da_L}{d\varphi} + \frac{d\Delta_k}{d\varphi} \lesseqgtr 0. \tag{A7}$$

From (17) to (18)

$$\frac{d\Delta_k}{da_U} = \varphi(1-s)\alpha\theta a_U^{\alpha-1} - (1-s)\lambda\beta a_U^{\beta-1} < 0,$$

$$\frac{d\Delta_k}{da_L} = -\varphi(1-s)\alpha\theta a_L^{\alpha-1} + (1-s)\lambda\beta a_L^{\beta-1} > 0, \tag{A8}$$

$$\frac{d\Delta_k}{d\varphi} = \theta\{s(1-\underline{a}^{\alpha}) + (1-s)(a_U^{\alpha} - a_L^{\alpha})\} > 0. \tag{A9}$$

Since $\beta > \alpha$, $\beta\lambda = \alpha\theta$ and $\varphi < 1$ and from (A5) and (A6), we have $\frac{da_U}{d\varphi} > 0$ and $\frac{da_L}{d\varphi} < 0$. Together, these derivatives imply that the indirect effect is always negative, i.e., $\frac{d\Delta_k}{da_U}\frac{da_U}{d\varphi} + \frac{d\Delta_k}{da_L}\frac{da_L}{d\varphi} < 0$, whereas the direct effect is positive, $\frac{d\Delta_k}{d\varphi} > 0$.

REFERENCES

Anderson, James and Eric van Wincoop. 2004. "Trade Costs," *Journal of Economic Literature*, XLII, 691–751.

Baldwin, Richard E., Rikard Forslid, Philippe Martin, Gianmarco I. P. Ottaviano, and Frederic Robert-Nicoud. 2003. *Economic Geography and Public Policy*, Princeton: Princeton University Press.

Baldwin, Richard E. and Toshihiro Okubo. 2006. "Heterogeneous firms, Agglomeration and Economic Geography: Spatial Selection and Sorting," *Journal of Economic Geography*, 6, 323–346.

Behrens, Kristian, Carl Gaigné, Gianmarco I.P. Ottaviano, and Jacques-Francois Thisse. 2006. "How Density Economics in International Transportation Link the Internal Geography of Trading Partners," *Journal of Urban Economics*, 60, 248–263.

Behrens, Kristian and Pierre M. Picard. 2011. "Transportation, Freight Rates, and Economic Geography," *Journal of International Economics*, 85, 280–291.

Borcherding, Thomas E. and Eugene Silberberg. 1978. "Shipping the Good Apples Out: The Alchian and Allen Theorem Reconsidered," *Journal of Political Economy*, 86, 131–138.

Bosker, Maarten and Harry Garretsen. 2010. "Trade Costs in Empirical New Economic Geography," *Papers in Regional Science*, 89, 485–511.

Forslid, Rikard and Toshihiro Okubo. 2012. "On the Development Strategy of Countries of Intermediate Size-An Analysis of Heterogeneous Firms in a Multi-region Framework," *European Economic Review*, 56, 747–756.

Fujita, Masahisa and Jacques-Francois Thisse. 2002. *Economics of Agglomeration*, Oxford: Oxford University Press.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.

Fukao, Kyoji, Kenta Ikeuchi, YoungGak Kim, and Hyeog Ug Kwon. 2011. "Do More Productive Firms Locate New Factories in More Productive Locations? An Empirical Analysis based on Panel Data from Japan's Census of Manufactures," RIETI Discussion Paper 11068.

Hummels, David. 1999. "Transport and International Trade," Report of the One Hundred and thirtieth Round Table on Transport Economics, OECD.

~~Hummels, David~~. 2007. "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21, 131–154.

Hummels, David and Alexandre Skiba. 2004. "Shipping the Good Apples Out: An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112, 1384–1402.

Martin, Philippe and Carol A. Rogers. 1995. "Industrial Location and Public Infrastructure," *Journal of International Economics*, 39, 335–351.

McCann, Philip. 2005. “Transport Costs and New Economic Geography,” *Journal of Economic Geography*, 5(3), 305–318.

———. 2001. “A Proof of the Relationship between Optimal Vehicle Size, Haulage Length and the Structure of Distance-Transport Costs,” *Transportation Research A*, 35, 671–693.

Melitz, Marc J. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71, 1695–1725.

Midelfart-Knarvik, Karen-Helene, Henry G. Overman, Stephen Redding, and Anthony J. Venables. 2002. “Integration and Industrial Specialisation in the European Union,” *Revue economique*, 53, 469–481.

Mori, Tomoya and Koji Nishikimi. 2002. “Economies of Transport Density and Industrial Agglomeration,” *Regional Science and Urban Economics*, 32, 167–200.

Okubo, Toshihiro, 2009. “Trade Liberalisation and Agglomeration with Firm Heterogeneity—Forward and Backward Linkages,” *Regional Science and Urban Economics*, 39, 530–541.

Okubo, Toshihiro and Eiichi Tomiura. 2013. “Regional Variations in Productivity Premium of Exporters: Evidence from Japanese Plant-level Data,” RIETI Discussion Paper 13005.

Okubo, Toshihiro, Pierre M. Picard, and Jacques-Francois Thisse. 2010. “The Spatial Selection of Heterogeneous Firms,” *Journal of International Economics*, 82, 230–237.

Ottaviano, Gianmarco I.P., Takatoshi Tabuchi, and Jacques-Francois Thisse. 2002. “Agglomeration and Trade Revisited,” *International Economic Review*, 43, 409–436.

Saito, Hisamitsu, Munisamy Gopinath, and JunJie Wu. 2011. “Heterogeneous Firms, Trade Liberalization and Agglomeration,” *Canadian Journal of Economics*, 44, 541–560.