

# Class Notes on New Economic Geography Models

Econ 8401-T.Holmes

(incomplete)

## 1 Overview

Last class we discussed various models of trade with increasing returns, e.g. those based on Dixit and Stiglitz models. These models took the standard international trade stance that goods are mobile while factors are fixed across locations. A second generation of models allows for factor mobility. This literature, the “New Economic Geography Literature” or NEG was initiated by Krugman (1991) and this work was cited extensively in the announcement for his 2008 Nobel Prize, titled “Trade and Trade and Geography – Economies of Scale, Differentiated Products and Transport Costs.” Often these are thought of as “regional models,” since factors of production are often regarded as mobile within a country’s borders but immobile across borders. Of course, factors also can move at the international level. In any case, for anyone interested in understanding the movement of goods and services across space, some basic knowledge of the NEG literature is useful.

The underpinning of the literature is a very simple idea. If we combine preference for variety and scale economies with transportation costs, there will be an incentive to concentrate economic activity in one place. By doing so, it is possible to produce a large number of goods at high volumes in the same place (and achieve high variety and exploit scale economies) yet not have to ship goods (and thereby avoid expenditures on transportation costs). Of course, this sounds like something people might have thought of before Krugman and of course this is true. What made the NEG literature big in the 1990s was the innovation of applying the Dixit-Stiglitz models to make it tractible to work out equilibria in market economies with these features. Regional scientists and economic geographers working earlier on these issues did not have this tool, so the analysis was either informal or based on social

planning problems.

I skip the original Dixit and Stiglitz CES treatment by Krugman. Instead I start with a version by Ottaviano, Tabuchi and Thisse (2002) or OTT which uses a Linear quadratic preference structure instead of the standard CES. You have already seen CES-based approaches. Linear quadratic approaches have attracted some attention (see also Melitz and Ottaviano (2008)), so this is something useful for you to see.

Next I go through an version that is based on BEJK structure and is something I am currently working on with Wen-Tai Hsu and Sanghoon Lee (two recent Minnesota Ph.D. graduates). The main interest in this paper is understanding the link between agglomeration and productivity.

Next I discuss an example empirical application: Redding and Sturm (2008)'s analysis of the reunification of Germany,

The last part of the lecture discusses the recent Arkolakis, Costinot, and Rodriguez-Clare (2010) paper on equivalence of various micro-based trade models for implications about aggregates. There is no NEG in this paper. However, the issues the paper raises are also relevant for NEG models, so this is a good place to talk about the paper.

## 2 The Linear Quadratic NEG model of OTT

### 2.1 Model

- Two regions,  $H$  and  $F$
- Two factors
  - Factor  $A$  evenly distributed
  - Factor  $L$  perfectly mobile
- Two sectors

- “Agriculture uses  $A$ , constant returns to scale one for one. The numeraire.
- “Manufacturing” uses  $L$ . Fixed cost of  $\phi$  to set up differentiated product, zero marginal cost
- Transportation cost of  $\tau$  in units of numeraire. Note: It is not an iceberg cost.
- Preferences
  - $q_0$  consumption of agricultural good
  - $q(i)$  consumption of differentiated good  $i$

$$U(q_0; q(i), i \in [0, N]) = \alpha \int_0^N q(i) di - \frac{\beta - \gamma}{2} \int_0^N [q(i)]^2 di - \frac{\gamma}{2} \left[ \int_0^N q(i) di \right]^2 + q_0$$

for

$$\alpha > 0$$

$$\beta > \gamma > 0$$

(Note that with linear demand there are no income effects. This makes this approach less likely to be useful for empirical applications compared to CES based approaches which have sensible income effects).

- Budget constraint given income  $y$  and endowment  $\bar{q}_0$ .

$$\int_0^N p(i)q(i)di + q_0 = y + \bar{q}_0$$

- Utility maximization for choice of  $q(i)$  yields linear demand functions

$$\alpha - (\beta - \gamma) q(i) - \gamma \int_0^N q(j) dj = p(i)$$

or

$$\begin{aligned}q(i) &= a - bp(i) + c \int_0^N [p(j) - p(i)]dj \\a &= \frac{\alpha}{\beta + (N - 1)\gamma} \\b &= \frac{1}{\beta + (N - 1)\gamma} \\c &= \frac{\gamma}{(\beta - \gamma)[\beta + (N - 1)\beta]}\end{aligned}$$

## 2.2 Equilibrium

- Mobile workers select location to maximize utility. Let  $\lambda$  be fraction that locate in  $F$ .
- Manufacturing firms decide whether to enter, where to locate and what price to set at each location, to maximize profit. (This setup implies zero profit.) Let  $n_H$  and  $n_F$  be the number (or measure) of firms at each location.
- Agricultural firms (competitive sector) maximize profits
- The usual market clearing conditions hold.

## 2.3 Derivation of Equilibrium

- Labor market clearing implies (Remember marginal cost is normalized to zero. Such a normalization never is used with CES structures, because price is a constant marginal cost) With linear demand

$$\begin{aligned}n_H &= \frac{\lambda L}{\phi} \\n_F &= (1 - \lambda)\frac{L}{\phi}\end{aligned}$$

- Let  $q_{ij}$  be sales of a firm at  $i$  at location  $j$ . Demand of a representative firm is

$$q_{HH}(p_{HH}) = a - (b + cN)p_{HH} + cP_H$$

$$q_{HF}(p_{HF}) = a - (b + cN)p_{HF} + cP_F$$

where

$$P_H = n_H p_{HH} + n_F p_{FH}$$

$$P_F = n_H p_{HF} + n_F p_{FF}$$

- Variable profit

$$\pi_H = p_{HH}q_{HH}(p_{HH}) + p_{HF}q_{HF}(p_{HF})$$

- It is clear equilibrium price depends upon entry  $(n_H, n_F)$ , so different from standard Dixit Stiglitz. The fact that markups are endogenous is considered a highlight of this model and it is the most attractive feature of the modelling structure.
- Given  $P_H$  and  $P_F$ , and  $N_H$  and  $N_L$ , solve linear equations to get profit maximizing  $p_{HH}$  and  $p_{HF}$ . Then the above equations are used to eliminate  $P_H$  and  $P_F$ . It is possible to derive linear pricing formulas,

$$p_{HH}^* = \frac{1}{2} \frac{2a + \tau c(1 - \lambda)N}{2b + cN}$$

$$p_{HF}^* = p_{FF}^* + \frac{\tau}{2}$$

- Make assumption that

$$\tau < \tau_{trade} \equiv \frac{2a\phi}{2b\phi + cL}$$

so that there is trade

- Equilibrium wage sets profit equal to zero

$$w_H^*(\lambda) = \frac{\pi_{HH}^* + \pi_{HF}^*}{\phi}$$

- Utility of a worker

$$V_H(\lambda) = S_H(\lambda) + w_H^*(\lambda) + \bar{q}_0$$

where  $S_H(\lambda)$  is consumer surplus associated with the equilibrium prices. It is strictly increasing and concave in  $\lambda$ .

- Equilibrium condition for worker location choice can be written as:

$$V_H(\lambda) = V_L(\lambda) \text{ if } \lambda \in (0, 1)$$

$$V_H(\lambda) \geq V_L(\lambda), \text{ if } \lambda = 1$$

$$V_H(\lambda) \leq V_L(\lambda) \text{ if } \lambda = 0$$

## 2.4 Results

- As is usual in this literature, there always exists a dispersed equilibrium ( $\lambda = .5$ ). This is obvious by symmetry.
- Add a stability definition (this is old fashioned since this is a static model, but commonly invoked in the literature).

*Def.* An allocation of workers  $\lambda \in (0, 1)$  is *stable* iff

$$\Delta(\lambda) = V_H(\lambda) - V_L(\lambda) = 0$$

$$\Delta'(\lambda) < 0$$

so “negative feedback.”

- Main result: There exists a cutoff  $\tau^*$  such that
  - $\tau > \tau^*$ , then the symmetric (dispersed) allocation is the only stable equilibrium.
  - $\tau < \tau^*$ , then agglomeration ( $\lambda = 1$  or  $\lambda = 0$ ) is the only stable equilibrium. (Called a "Black Hole" in the literature).
- Intuitive comparative static results can be derived analytically
  - $\tau^*$  increases with  $\phi$
  - $\tau^*$  increases with more product differentiation (i.e. falls with  $\gamma$ ).
- Efficiency. Consider a social planner maximizing total surplus.
  - There exists a  $\tau^0 < \tau^*$ , so that  $\tau < \tau^0$ , agglomerate,  $\tau > \tau^0$ , dispersion
  - So in region  $\tau \in (\tau^0, \tau^*)$  have "excess agglomeration"

### 3 A BEJK NEG Model

#### 3.1 Introduction (abbreviated)

This is work in progress that is joint with Wen-Tai Hsu and Sanghoon Lee. (These notes are actually just me cutting and pasting pieces of the paper. Sorry about that!)

Let me provide some background motivation. As has been noted in this class, models of firm heterogeneity that explicitly take into account exit by low productivity firms have played a prominent role in the international trade literature, with Melitz-based models being the market leader as compared to Bernard, Eaton, Jensen Kortum (2003) (BEJK). Recently, these ideas have been applied to models of regions rather than countries and all of these new papers follow a Melitz approach.

This paper develops a regional analysis that incorporates heterogeneity in productivity, but in contrast to other papers, it follows the BEJK approach. To understand what we

do, it is first necessary to highlight two essential differences between the BEJK and the Melitz approaches. First, in BEJK, competition between firms is *head-to-head*. There is more than one potential producer of any given product and the different producers engage in Bertrand competition, market by market. In contrast, in the Melitz approach each firm has a monopoly over a particular differentiated product, as in Dixit and Stiglitz. Second, in BEJK, firms draw their productivity distribution from a distribution with a *fat right tail*. (To be more specific, firms draw from the Frechet distribution, but more on that later.) In contrast, in the Melitz approach, it is not essential that productivity be drawn from a fat tail.

This paper takes the two essential ingredients of the BEJK approach: head-to-head competition and fat-tailed productivity draws. It adds to this: (1) labor mobility, and (2) a model of freely-mobile entrepreneurial activity that endogenizes the distribution of productivity across locations. The results we get can be divided into three parts.

First, we succeed in creating a regional version of BEJK. Second, we show that imposing equilibrium conditions in the BEJK structure for mobile labor and mobile entrepreneurship has content. In particular, if there is agglomeration, then plants in large locations tend to be more productive than plants in small locations. Without imposing these equilibrium location choice conditions, this doesn't necessarily hold.

Third, we take the micro productivity data generated by our model and analyze it using approaches taken in three well known papers: Syverson (2004), CDGPR, and Hsieh and Klenow (2009). The analysis of the productivity data in our model yields very different conclusions from what is found in these three papers.

## 3.2 Model

There are two locations,  $i = 1, 2$ , that are ex ante identical. In the equilibrium of the model, it may happen that one location attracts more people than the other. We label things in such cases so that location 1 is the “big city” and location 2 is the “small city.”



All agents have the same preferences for a composite good  $Q$  and land  $L$ ; these preferences are represented by the utility function

$$U(Q, L) = Q^\beta L^{1-\beta}.$$

The composite is an aggregation of differentiated goods indexed by  $j$  on the unit interval. It follows the standard CES form,

$$Q = \left( \int_0^1 (q(j))^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}},$$

where  $\sigma$  is the elasticity of substitution.

The land supply in each location is the same and equals  $\bar{L}$ .

There is a measure  $\bar{H}$  individuals in the economy. Individuals first choose whether to live and work in city 1 or live and work in city 2. Next, they choose whether to be an entrepreneur or to be employed as a worker. Let  $N_i$  be the number of individuals choosing to be a worker in city  $i$  and let  $M_i$  be the number of entrepreneurs. The resource constraint implies that

$$N_1 + M_1 + N_2 + M_2 = \bar{H}.$$

It will be convenient to work with fractions. Define these by

$$n_i = \frac{N_i}{\bar{H}}, \quad m_i = \frac{M_i}{\bar{H}}.$$

We now explain the process through which firms are created and productivities are determined, beginning with the arrival of  $M_i$  entrepreneurs at location  $i$ . Each entrepreneur picks a product  $j \in [0, 1]$  to attempt to enter. Let  $S_i(j)$  be the density of entrepreneurs attempting to enter product  $j$  (the number of startups for this product). All entrepreneurs

arriving at  $i$  pick some industry; i.e.,

$$M_i = \int_0^1 S_i(j) dj.$$

Each entrepreneur entering an industry obtains a plant. The productivity of a plant has two components that enter multiplicatively. First, there is a term  $A_i$  that is constant across all plants at location  $i$  and depends on the amount of labor market spillovers. If there are  $N_i$  workers located at  $i$ , then

$$A_i = N_i^\zeta.$$

The parameter  $\zeta$  governs the significance of agglomeration spillovers. In particular, if  $\zeta = 0$ , then  $A_i = 1$  and there are no spillovers. Second, there is a random term  $y$  to productivity that depends upon the entrepreneur's luck. An entrepreneur at location  $i$  with productivity  $y$  for a particular good  $j$  can produce  $A_i y$  units of good  $j$ , per unit of labor procured at location  $i$ . For the purposes of this lecture, assume entrepreneurs draw from Frechet

$$G(y) = e^{-y^{-\theta}}$$

(In the paper we assume the underlying draws come from the more general class of fat-tailed distributions that include the Pareto. Taking extreme values of fat-tailed distributions maps into Frechet)

We assume an iceberg transportation cost  $\tau$  between the two locations. To deliver one unit of any differentiated good  $j$  to a different location,  $\tau \geq 1$  units must be shipped.

There are three stages in the model. In stage 1, individuals choose where to live and what job to hold. We assume that the  $\bar{L}$  units of land at location  $i$  are owned by landowners who live at  $i$  but do not work. They have the same utility function as the other individuals.

In stage 2,  $M_i$  entrepreneurs at location  $i$  allocate themselves across the product space  $j \in [0, 1]$  so that  $S_i(j)$  is the density choosing good  $j$ . Each of the  $S_i(j)$  entering product  $j$

at  $i$  obtains a single plant, with one draw of the random productivity term  $y$ . We impose as an equilibrium condition that the returns to entering each product  $j$  at a location are equalized. We ignore integer constraints

In stage 3, the  $S_1(j)$  plants at location 1 and the  $S_2(j)$  plants at location 2 engage in Bertrand price competition for the product  $j$  market at each location. At the same time there is market clearing in the labor markets and land markets.

### 3.3 Equilibrium for Fixed Location and Job Choices

We first show that entrepreneurs that enter a location spread out across the various products so  $S_i(j) = M_i$  entrepreneurs enter each product. Given the properties of the Frechet, it follows that the distribution of the highest productivity firm for a given product at  $i$  is Frechet

$$F_i(z) = e^{-T_i z^{-\theta}}, \quad (1)$$

with scaling parameter

$$T_i = n_i^{\theta\zeta} m_i. \quad (2)$$

Fixing  $n_i$  and  $m_i$ , the model maps into the international trade model of BEJK that we discussed earlier. We can then use their results.

In particular, define  $\Phi_n$  by

$$\Phi_n = \sum_{i=1}^2 T_i (w_i \tau_{ni})^{-\theta}. \quad (3)$$

This parameter  $\Phi_n = \sum_{i=1}^2 T_i (w_i \tau_{ni})^{-\theta}$  distills the parameters of productivity distributions, wages, and the trade cost into one single term governing the cost and price distributions.

The price distribution at location  $n$  is given by

$$K_n(p) = 1 - [1 + \Phi_n (1 - \mu^{-\theta}) p^\theta] e^{-\Phi_n p^\theta}.$$

The BEJK's analytical results that are useful for our paper are listed as follows.

**BEJK Result 1** The probability that location  $i$  provides a good at the lowest price in location  $n$  is

$$\pi_{ni} = \frac{T_i(w_i\tau_{ni})^{-\theta}}{\sum_{k=1}^2 T_k(w_k\tau_{nk})^{-\theta}} = \frac{T_i(w_i\tau_{ni})^{-\theta}}{\Phi_n}.$$

**BEJK Result 2** In any location  $n$ , the probability of buying a good with price lower than  $p$  is independent from where the good is purchased from. Letting  $X_{ni}$  be the total expenditure of location  $n$  on the goods from  $i$ , and  $X_n$  be the total expenditure, we have

$$X_{ni} = \pi_{ni}X_n.$$

**BEJK Result 3** Assume that  $\theta + 1 > \sigma$ . Let  $\Gamma$  denote the gamma function; the price index is

$$\begin{aligned} P_n &= \left[ \frac{1 + \theta - \sigma + (\sigma - 1)\mu^{-\theta}}{1 + \theta - \sigma} \Gamma\left(\frac{\theta + 1 - \sigma}{\theta}\right) \right]^{\frac{1}{1-\sigma}} \Phi_n^{-\frac{1}{\theta}} \\ &\equiv \gamma \Phi_n^{-\frac{1}{\theta}}. \end{aligned} \tag{4}$$

**BEJK Result 4** A fraction  $\theta/(1 + \theta)$  of revenue goes to variable cost.

Finally, we note that since trade between locations 1 and 2 is balanced, total expenditure  $X_n$  on goods at  $n$  is the same as total revenues of plants located at  $n$ .

### 3.4 Equilibrium Location and Job Choices

We begin by looking at job choice. Let  $w_i$  be the wage at  $i$  and  $v_i$  the expected profit of being an entrepreneur. Since the share of revenues going to workers and entrepreneurs is  $\theta/(1 + \theta)$  and  $1/(1 + \theta)$ , and since there are  $N_i$  and  $M_i$  of each, the returns to each job equal

$$w_i = \frac{\theta}{1 + \theta} \frac{X_i}{N_i}, \tag{5}$$

$$v_i = \frac{1}{1 + \theta} \frac{X_i}{M_i}. \tag{6}$$

Free mobility between the two jobs implies that  $w_i = v_i$ , which implies a constant ratio between workers and entrepreneurs in each location  $i$ , that is,

$$\frac{N_i}{M_i} = \frac{n_i}{m_i} = \theta. \quad (7)$$

We can use this result to rewrite the Fréchet scale parameter from equation (2) for the distribution of the maximum productivity plant at  $i$  as

$$T_i = n_i^{\theta\zeta} m_i = \theta^{\theta\zeta} m_i^{\theta\zeta+1}. \quad (8)$$

Since the utility is Cobb-Douglas in goods and land, the indirect utility for an individual choosing to locate at  $i$  given composite goods price  $P_i$ , land rental  $R_i$ , and wage  $w_i$  equals

$$U_i = P_i^{-\beta} R_i^{-(1-\beta)} w_i. \quad (9)$$

Since the local land-owners have the same Cobb-Douglas, expenditures on the  $\bar{L}$  units of land at  $i$  must equal the land expenditure share times location  $i$  income,

$$R_i \bar{L} = (1 - \beta)[R_i \bar{L} + (N_i + M_i)w_i].$$

solving for the rent  $R_i$  and substituting this into utility (9) yields

$$U_i = k \left( \frac{w_i}{P_i} \right)^\beta \left( \frac{1}{N_i + M_i} \right)^{1-\beta}, \quad (10)$$

where  $k$  is constant across the two locations.

Let  $f_i$  denote the fraction of individuals choosing to locate at  $i$ ,  $f_i \equiv (N_i + M_i) / \bar{H}$ . For a fixed value of the population shares  $f_1$  (and hence fixed  $f_2 = 1 - f_1$ ), we can impose the equilibrium condition for job choice (7) and we can solve for the  $w_1, w_2, P_1$  and  $P_2$  that are consistent with equilibrium in the output market. Plug these into (10) and let  $\tilde{U}_i(f_1)$  be

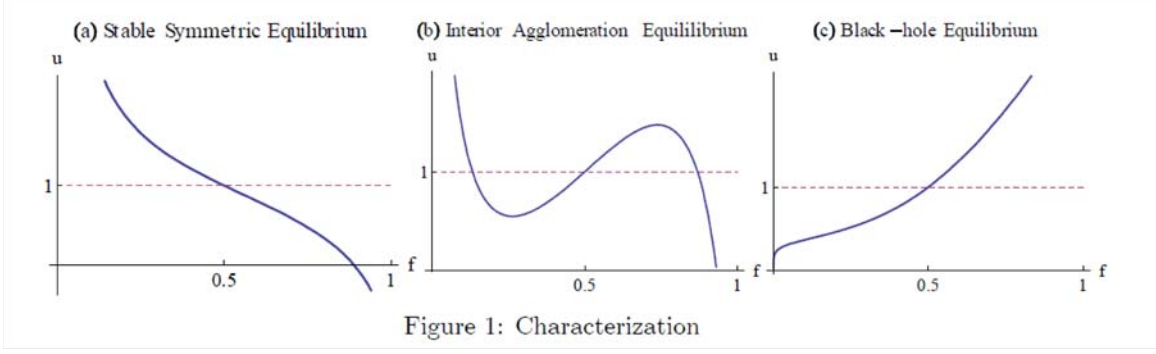


Figure 1:

the utility conditioned on  $f_1$ . Let  $u(f_1) = \tilde{U}_1(f_1)/\tilde{U}_2(f_1)$  be the ratio of utilities. For an interior value  $f_1 \in (0, 1)$  to be an equilibrium of location choice, it must be that  $u(f_1) = 1$ .

As is standard in this literature, a *symmetric equilibrium* always exists at  $f_1^* = 0.5$ , where half of the individuals go to each location. Given symmetry, we can restrict attention to the range  $f_1 \geq \frac{1}{2}$  where location 1 is weakly larger than location 2. An equilibrium with  $f_1^* > 0.5$  is called an *agglomeration equilibrium*. If  $u(1) \geq 1$ , then  $f_1^* = 1$  is an equilibrium where everyone goes to location 1. Call this a *black-hole equilibrium*. Define an interior equilibrium  $f_1^* \in [\frac{1}{2}, 1)$  as *stable* if  $du/df_1 < 0$  at  $f_1^*$  and *unstable* if  $du/df_1 > 0$ . The Proposition below shows that depending on the parameters, there are three possibilities for how things can look. Figure 1 illustrates the three cases, showing how the utility ratio  $u$  depends upon the share  $f_1$ .

**Proposition 1** *Define two thresholds for  $\beta$  as a function of other model parameters,*

$$\hat{\beta}_L \equiv \frac{1}{\left(\zeta + 1 + \frac{1}{\theta}\right)} \frac{\tau^\theta(1 + 2\theta) + 1}{\tau^\theta(1 + 2\theta) - 1}, \quad \hat{\beta}_H = \frac{1 + \theta}{1 + \theta + \theta\zeta} .$$

(i) *The parameter space  $\{(\beta, \tau, \theta, \zeta) : \beta \in (0, 1), \tau > 1, \theta > 1, \zeta \geq 0\}$  can be partitioned into the the following three subspaces each of which is associated with a distinct characterization of equilibria.*

(a) *When  $\beta \leq \hat{\beta}_L$ , the symmetric equilibrium is stable and the only equilibrium.*

(b) When  $\beta > \hat{\beta}_L$ , the symmetric equilibrium is unstable, and there exists a unique agglomeration equilibrium that is stable. If  $\beta \in (\hat{\beta}_L, \hat{\beta}_H)$  the agglomeration equilibrium is interior,  $f_1^* \in (0.5, 1)$ , and if  $\beta \geq \hat{\beta}_H$ , the agglomeration is a black-hole,  $f_1^* = 1$ .

(ii) If there is an interior agglomeration equilibrium  $f_1^* \in (0.5, 1)$ , then  $f_1^*$  increases in  $\beta$ ,  $\tau$ , and  $\zeta$ .

(iii) For large enough  $\tau$ , if there is an interior agglomeration equilibrium  $f_1^* \in (0.5, 1)$ , then  $f_1^*$  increases as  $\theta$  decreases.

As shown in Helpman (1998), agglomeration increases with the weight  $\beta$  placed on goods consumption relative to land consumption and with the transportation cost  $\tau$ . That agglomeration increases with the knowledge spillover parameter  $\zeta$  is a standard finding.

The new result here is the way selection is highlighted as a force of agglomeration. The shape parameter  $\theta$  of the Frechet determines the dispersion of productivity. The lower is  $\theta$ , the more disperse is productivity, and therefore the more important is selection. When  $\tau$  is big, making selection more important by decreasing  $\theta$  leads to more agglomeration. Selection over productivity replaces the role of preference for variety (through lower  $\sigma$ ) found in standard NEG models as an agglomerating force.

### 3.5 Productivity and Agglomeration

This section analyzes the distribution of productivities of surviving plants at the two locations. The first part derives the implications of the equilibrium conditions of worker and entrepreneurial choice. The conditions imply that the distribution of productivity is higher in the large city compared to the small city. The second part compares the observed productivity distributions to those generated in a benchmark model where selection over productivity is shut down. The analysis shows it is impossible to distinguish these two cases

with data on observed productivity distributions. The analysis also shows that it matters to distinguish the two cases—the two cases differ fundamentally in their underlying economics.

### 3.5.1 The Productivity Distribution is Higher in the Large City

We begin by deriving two equations from the equilibrium conditions. Recall that  $\Phi_i$  defined in (3) is the summary statistic that pins down the distribution of prices at location  $i$ , gathering together all the various forces including the productivity distributions at each location. Define  $\phi \equiv \Phi_1/\Phi_2$  as the ratio of this key statistic between the large and small cities and analogously define  $f \equiv f_1/f_2$  to be the ratio of population shares. We derive two conditions linking  $\phi$  and  $f$ . One equation uses indifference between wage work and entrepreneurship; the other uses indifference about where to live.

Using BEJK Results 1 and 2, expenditure by location 1 on goods from location 2 equals  $X_{12} = X_1 T_2 (w_2 \tau)^{-\theta} / \Phi_1$ . Using the analogous expression for  $X_{21}$  and the market clearing condition  $X_{12} = X_{21}$  implies

$$\frac{X_1}{X_2} = \frac{T_1 w_1^{-\theta} \Phi_1}{T_2 w_2^{-\theta} \Phi_2} = \frac{m_1^{\theta\zeta+1}}{m_2^{\theta\zeta+1}} w^{-\theta} \phi,$$

where we substitute in equation (8)  $T_i = \theta^{\theta\zeta} m_i^{\theta\zeta+1}$  for the productivity distribution scaling parameter and we let  $w \equiv w_1/w_2$  be the wage ratio. The equilibrium job choice condition (7) implies  $n_1/n_2 = m_1/m_2 = f$ . Also,  $wf = X_1/X_2$  (equation (5)). Plugging these in gives

$$w^{1+\theta} = f^{\theta\zeta} \phi. \tag{11}$$

Next we use the definition (3) of  $\Phi_i$  to obtain

$$\phi = \frac{\Phi_1}{\Phi_2} = \frac{T_1 w_1^{-\theta} + T_2 w_2^{-\theta} \tau^{-\theta}}{T_1 w_1^{-\theta} \tau^{-\theta} + T_2 w_2^{-\theta}} = \frac{f^{\theta\zeta+1} w^{-\theta} + \tau^{-\theta}}{f^{\theta\zeta+1} w^{-\theta} \tau^{-\theta} + 1},$$

where again we use the job choice equilibrium condition to substitute in  $f$  for  $m_1/m_2$ .



Solving this expression for  $w$  and substituting into (11) yields our first equation linking  $f$  and  $\phi$ ,

$$f^{\theta\zeta+\theta+1} = \left( \frac{\phi - \tau^{-\theta}}{1 - \tau^{-\theta}\phi} \right)^{1+\theta} \phi^\theta. \quad (12)$$

To derive the second equation in  $f$  and  $\phi$ , we use the indifference condition from the individual's choice of where to live. This implies that the ratio of utilities  $u(f_1) = 1$  for an interior equilibrium,  $f_1^* < 1$ , i.e.,

$$\begin{aligned} 1 &= \frac{U_1}{U_2} = \frac{\left(\frac{P_1}{w_1}\right)^{-\beta} (N_1 + M_1)^{-(1-\beta)}}{\left(\frac{P_2}{w_2}\right)^{-\beta} (N_2 + M_2)^{-(1-\beta)}} \\ &= w^\beta \phi^{\frac{\beta}{\theta}} f^{-(1-\beta)}. \end{aligned}$$

Using (11) to substitute in for  $w$ , we obtain our second equation

$$f^{(1-\beta)(1+\theta)-\beta\zeta\theta} = \phi^{\beta(1+2\theta)/\theta}. \quad (13)$$

With the derivation of conditions (12) and (13) complete, we can now analyze productivity and selection. If a plant for a particular product  $j$  at location  $i$  survives, it is necessarily the most efficient plant for product  $j$  at  $i$ . As derived in Section 3, the distribution of the most efficient plant at  $i$  for a given  $j$  is Fréchet, with density  $g_i(z) = T_i \theta z^{-\theta-1} e^{-T_i z^{-\theta}}$ , where  $T_i$  and  $\theta$  are the scaling and shape parameters. In addition, for the plant to survive, its cost must be lower than what it would cost the most efficient plant at the other location to export. Recalling our notation that  $Z_{1i}$  is the productivity of the most efficient plant at  $i$ , then the most efficient firm at location 1 survives if and only if

$$\frac{w_1}{Z_{11}} < \frac{w_2 \tau}{Z_{12}}.$$

We can use this to calculate the productivity distribution of the most efficient plant at 1,

conditioned on beating the competition at 2,

$$\begin{aligned}
\frac{\Pr(Z_{11} \leq z, Z_{11} > \frac{w_1}{w_2\tau} Z_{12})}{\Pr(Z_{11} > \frac{w_1}{w_2\tau} Z_{12})} &= \frac{1}{\pi_{11}} \Pr(Z_{11} \leq z, Z_{11} > \frac{w_1}{w_2\tau} Z_{12}), \\
&= \frac{1}{\pi_{11}} \int_0^z \int_0^{\frac{w_2\tau}{w_1} z_{11}} g_1(z_{11}) g_2(z_{12}) dz_{12} dz_{11}, \\
&= e^{-(w_1^\theta \Phi_1) z^{-\theta}}.
\end{aligned}$$

Thus the productivity distribution of surviving plants at 1 is Fréchet with shape parameter  $\theta$ , and scaling parameter equal to

$$\begin{aligned}
\hat{T}_1 &= w_1^\theta \Phi_1 \\
&= T_1 + T_2 \left( \frac{w_1\tau}{w_2} \right)^{-\theta}.
\end{aligned}$$

The selection induced by competition with the other location increases the scaling from  $T_1$  to a higher level  $\hat{T}_1$ ; i.e., it shifts the distribution to the right. Similarly, the productivity distribution of survivors in location 2 has scale parameter  $\hat{T}_2 = w_2^\theta \Phi_2$ , but otherwise has the same shape parameter  $\theta$ . Therefore the surviving plants at location 1 have a higher distribution than the survivors at 2 (in the sense of first-order stochastic dominance) if  $\hat{T}_1 > \hat{T}_2$ . Our result is

**Proposition 2** *Suppose there is an agglomeration at location 1 (the large city), i.e.,  $f > 1$ .*

(i) *The productivity distributions of survivors are Fréchet at both locations with the same shape parameter  $\theta$ , but the scaling parameter  $\hat{T}_1 = w_1^\theta \Phi_1$  at location 1 is strictly higher than the scaling parameter  $\hat{T}_2 = w_2^\theta \Phi_2$  at the large city.*

(ii) *The ratio of the mean productivities of the survivors equals*

$$\frac{E[Z_1 | \text{survive}]}{E[Z_2 | \text{survive}]} = \frac{\frac{w_1}{P_1}}{\frac{w_2}{P_2}} = f^{\frac{1-\beta}{\beta}}$$

which is strictly greater than one. The variance of the productivity distribution of survivors is higher at the large city.

(iii) The mean of the logarithm of survivors' productivities is higher in the large city, while the variance of log productivity is constant across the two locations.

*Proof of (i).* We need to show that  $\widehat{T}_1/\widehat{T}_2 = w^\theta \phi > 1$ . This equals

$$\begin{aligned} w^\theta \phi &= f\left(\frac{\theta c}{1+\theta}\right)^\theta \phi^{\frac{\theta}{1+\theta}} \phi \\ &= f^{\frac{(1-\beta)\theta}{\beta}} > 1 \end{aligned}$$

To obtain the first line, we substitute in for  $w$  using (11). To obtain the second line, we solve out (13) for  $\phi$  in terms of  $f$  and substitute in. *Proof of (ii).* For the Fréchet with scaling parameter  $T$ , the mean and variance equal  $T^{1/\theta}\Gamma((\theta-1)/\theta)$  and  $T^{2/\theta}[\Gamma((\theta-2)/\theta) - \Gamma^2((\theta-1)/\theta)]$ . Hence,

$$\frac{E[Z_1|\text{survive}]}{E[Z_2|\text{survive}]} = \left(\frac{\widehat{T}_1}{\widehat{T}_2}\right)^{\frac{1}{\theta}} = w\phi^{\frac{1}{\theta}}.$$

That this equals  $f^{\frac{(1-\beta)\theta}{\beta}}$  follows from the proof of part (i). That it equals the ratio of the real wages follows from the definition of  $\phi$  and BEJK Result 3. That variance is higher in the big city follows from the formula for variance.

### 3.5.2 Shutting Down Selection: Does it Look Different and Does it Matter?

We highlight the role of selection in our model by contrasting a version of model in which spillovers are shut down with a *benchmark model* in which selection is completely shut down. While the economics of the benchmark model is different, the two models look the same in terms of observed productivity distributions.

The key assumption of the benchmark model is that there is a monopoly entrepreneur for each product  $j \in [0, 1]$  rather than free entry and head-to-head competition. Each of the unit measure of entrepreneurs chooses where to locate. After the location decision is

fixed, each entrepreneur draws a random productivity term  $y$  from the Frechet with scale parameter  $T^b$  and shape parameter  $\theta^b$ . (We use the superscript “ $b$ ” to denote parameters of the benchmark model.) There are a unit measure of workers that also choose where to live. Let  $n_i$  be the fraction of workers locating at  $i$  and  $m_i$  be the fraction of entrepreneurs,  $n_1 + n_2 = 1$ ,  $m_1 + m_2 = 1$ . The productivity of an entrepreneur at  $i$  drawing  $y$  equals  $A_i y$ , where  $A_i = n_i^{\zeta^b}$ , for spillover parameter  $\zeta^b$ . To complete the model, let  $\sigma^b$  be the elasticity of substitution for the composite utility function, let  $\tau^b$  be the transportation cost, and  $\beta^b$  be the goods share of utility.

The benchmark model is exactly the Helpman NEG model with two exceptions. The first exception is a fixed set of monopolist over a fixed variety of products, rather than the usual free entry with fixed costs. This makes no difference for what we do. The second exception is that productivity is random rather than deterministic. This makes no difference in deriving the equilibrium, as the integral involving the expectation of productivity draws factors out. (It does make a difference in the data as it will generate a Frechet distribution of productivities.) In the equilibrium of the model, the fraction of entrepreneurs locating at  $i$  is the same as the fraction of workers,  $m_i = n_i = f_i$ , so let  $f = f_1/f_2$  be the population ratio as before. Fixing  $f$ , the model is standard Dixit-Stlitz, such that a firm at  $i$  drawing  $y$  sets price equal to a constant markup over cost,  $p_i(y) = \frac{\sigma^b w_i}{\sigma^b - 1} / A_i y$ .

We compare the benchmark model to a version of our model in which knowledge spillovers are shut down. We will refer to this as the *selection model*, with parameters  $(\beta^s, \tau^s, \theta^s)$ , and spillover  $\zeta^s = 0$ . First, we draw a connection between the way equilibria look in the selection and benchmark models.

**Proposition 3** *Fix the parameters  $(\beta^s, \tau^s, \theta^s)$  of our selection model, with  $\zeta^s = 0$ . Assume the conditions in Proposition 1 for an interior agglomeration equilibrium hold and let  $f^*$  be the equilibrium population ratio. Define a parameter set for the benchmark model with the exact same values for transportation costs and goods consumption share,  $\tau^b = \tau^s = \tau$ ,  $\beta^b = \beta^s = \beta$ . Let  $\theta^b = \theta^s$ , and  $\sigma^b = \theta^s + 1$ , and set the knowledge spillover parameter equal*

to

$$\zeta^b = \frac{1 - \beta}{\beta}.$$

- (i) The equilibrium population ratio  $f^*$  is the same in both the selection and the benchmark models.
- (ii) The distribution of productivities of producing plants at each location is that same in both the selection and benchmark models.

We will discuss the proof in class. This relates to the equivalence of Dixit-Stiglitz and the BEJK model pointed out by Arkolakis, Costinot, and Rodriguez-Clare (2010)

While these models look similar in the data they generate, the underlying economics are different and optimal policy is different. We illustrate this by considering the welfare impacts of a *zoning policy* that permits production only at location 1, i.e., suppose  $f_1 = 1$  and  $f_2 = 0$  are mandated by policy. (See Rossi-Hansberg (2004), for example.) The differing impacts of zoning in the two models are put in sharp contrast by an analysis of the limiting case where there is no transportation cost.

**Proposition 4** *Assume  $\tau = 1$ . The equilibrium outcome is the same in both models, equal dispersion across the two locations,  $f_1 = f_2 = \frac{1}{2}$ . However, the welfare effect of the zoning policy differs across the two models. The zoning policy strictly decreases aggregate utility in the selection model for any value of the model parameters. In the benchmark model, zoning increases aggregate utility if and only if*

$$\zeta > \frac{1 - \beta}{\beta}.$$

**Proof.** That the zoning policy reduces aggregate utility in the selection model is immediate. In the benchmark model, if  $f_1$  and  $f_2$  are the population fractions then aggregate utility equals

$$U^b = \left(f_1^{1+\zeta}\right)^\beta L^{1-\beta} + \left(f_2^{1+\zeta}\right)^\beta L^{1-\beta}$$

To see this, observe that if  $f_i$  locate at  $i$ , then goods production there equals  $f_i^{1+\zeta}$ , taking account of the knowledge spillover. Straightforward calculations show that if  $\zeta > \frac{1-\beta}{\beta}$ , this is maximized at  $f_1 = 1$  and  $f_2 = 0$ , while if  $\zeta < \frac{1-\beta}{\beta}$ , this is maximized at  $f_1 = f_2 = 0.5$ . ■

Let us now bring all the ideas of this subsection together. Imagine we have a data generating process that is either the benchmark model or the selection model, we don't know which. Suppose we observe transportation cost and it is initially positive,  $\tau > 1$ , and we have an agglomeration at location 1. We observe the productivity distribution in both cities and see that it is higher at location 1, but, as we know from Proposition 3, access to the micro data doesn't help distinguish between the two models. Suppose transportation costs fall to zero, i.e.,  $\tau_{new} = 1$ . Both models have the same prediction for equilibrium comparative statics: agglomeration disappears and we move to an equal split of population. However, the policy implications are very different in the two models. If selection is the true source of the productivity gains in the large city, then a zoning policy at the new transportation cost strictly decreases welfare. In contrast, if the source is knowledge spillovers, the policy is welfare neutral. (From Proposition 3,  $\zeta = \frac{1-\beta}{\beta}$ , and this value is the borderline case for welfare in Proposition 4.)

We can use this analysis to examine an idea in Combes, Duranton, Gobillon, Puga, and Roux (2009) (hereafter CDGPR) to distinguish selection from spillovers. This is what one gets when using Melitz-style models with no head-to-head competition. The idea is to look for truncation on the left as evidence of selection and any smooth overall rightward shift as evidence of spillovers. In their empirical analysis, CDGPR find little evidence of any kind of increased truncation in large cities. Rather they find productivity distributions shift to the right in a relatively smooth way. They conclude that spillovers must be what drives productivity gains. However, in our model, the pattern they document in the data is equally consistent with all productivity gains being due to selection and none being due to spillovers. As shown in Proposition 3, "Selection with No Spillover" and "Spillover with No Selection" generate the same kind of productivity data, shifting the distribution to the

right in a smooth way in large cities.

Proposition 3 is related to Arkolakis, Costinot, and Rodriguez-Clare (2010), that argues that adding firm heterogeneity in either a Melitz or BEJK fashion does not add anything new in terms of aggregate impacts beyond what is already in the Dixit-Stiglitz symmetric firm model. They show these different models of international trade are equivalent at the aggregate level. The equivalence result between Dixit-Stiglitz and BEJK appears here as well. As noted above, the benchmark model is essentially Helpman (1998) and Helpman is Dixit-Stiglitz. The benchmark model looks the same (with the same comparative statics from a change in transportation cost  $\tau$ ) as the selection model, which is BEJK in its underlying moving parts. While the connections noted by Arkolakis, Costinot, and Rodriguez-Clare (2010) show up here as well, the bottom line point that we get in our regional model with mobile labor and entrepreneurship is quite different from what they get in their trade model with fixed factors. The welfare effects of policies that impact the movement of factors of production do depend upon the underlying model. Whether it is the benchmark or the selection model matters.

## 4 Redding and Sturm (2008)

Redding and Sturm apply NEG theory to understand how cutting off market access—through the division of postwar Germany—impacts where people live. The paper does two things. First, through a dif-dif exercise the paper shows the impacts are consistent with the qualitative implications of NEG theory. Second, the paper calibrates an NEG model and it argues that the quantitative implications of the theory work well. Rather than say discuss the qualitative results in the notes, we will just put the pdf of the paper up in the overheads in class and scan through some of the tables and maps. Here we will just discuss a little bit about the model. I should say that the quantitative aspects of this kind of research is where the interesting work is going forward. In general, simple two location models NEG models

have been worked out and people are tired of them.

The model is based on Helpman (1998). (Some of these equations come from the appendix for the Redding and Sturm paper at the AER website.)

- There are  $L$  mobile consumers each with a unit of time.
- $U_c = (C_c^M)^\mu (C_c^H)^{1-\mu}$ ,  $0 < \mu < 1$ , at city  $c$ .
- tradeable consumption  $C_c^M$  standard CES,  $\sigma$  is elasticity of substitution
- $C_c^H$  is nontradeable amenity, exogenously distributed across locations.
- $T_{ic} > 1$  iceberg cost to ship from  $i$  to  $c$ .
- Fixed cost of  $F$  (labor units) to set up a new product, constant marginal marinal cost of one unit of labor. So to produce  $x$  units requires  $l = F + x$  units of labor

#### 4.1 Equilibrium conditions on the Demand Side

The price index is at city  $c$  is

$$P_c^M = \left[ \sum_i n_i (p_i T_{ic})^{1-\sigma} \right]^{1/(1-\sigma)}$$

where  $n_i$  is variety at location  $i$

Define consumer market access  $CMA_c$  by

$$\begin{aligned} P_c^M &= [CMA_c]^{1/(1-\sigma)} \\ CMA_c &\equiv \sum_i n_i (p_i T_{ic})^{1-\sigma} \end{aligned}$$

Now expenditure is (where  $E_c$  is total expenditure at  $c$ ),

$$\mu E_c = P_c^M Q_c$$



From Shepard's Lemma (an envelope theorem),

$$\frac{\partial [P_c^M]}{\partial p_i} Q_c = n_c x_{ic}$$

Or

$$\frac{1}{1-\sigma} \left[ \sum_i n_i (p_i T_{ic})^{1-\sigma} \right]^{1/(1-\sigma)-1} (1-\sigma) n_i (p_i T_{ic})^{-\sigma} T_{ic} Q_c = n_c x_{ic}$$

But can write the above as

$$\left[ \sum_i n_i (p_i T_{ic})^{1-\sigma} \right]^{-1} n_i (p_i T_{ic})^{-\sigma} T_{ic} P_c^M Q_c = n_c x_{ic}$$

Or

$$\left[ \left[ \sum_i n_i (p_i T_{ic})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \right]^{(\sigma-1)} n_i (p_i T_{ic})^{-\sigma} T_{ic} \mu E_c = n_c x_{ic}$$

or

$$x_{ic} = p_i^{-\sigma} (T_{ic})^{1-\sigma} (\mu E_c) (P_c^M)^{\sigma-1}$$

The bottom line equation above is equilibrium demand for a tradeable variety from city  $c$ .

There is an inelastic supply of the non-tradeable amenity  $H_c$ . It must be the case that the price of the amenity at  $c$  must equal

$$P_c^H = \frac{(1-\mu) E_c}{H_c}.$$

Total expenditure is the sum of labor and amenity expenditure,

$$E_c = w_c L_c + (1-\mu) E_c = \frac{w_c L_c}{\mu}$$

Note this implies that  $\mu E_c = w_c L_c$ . Let's plug this in above for the price of amenities,

$$P_c^H = \frac{(1-\mu) E_c}{H_c} = \frac{(1-\mu) w_c L_c}{H_c \mu}$$

## 4.2 Equilibrium Conditions on the Supply Side

As usual the *free-on-board* price (buyer pays transportation costs) is

$$p_i = \frac{\sigma}{\sigma - 1} w_i$$

Get usual fixed plant size in free entry zero profit equilibrium

$$\bar{x} = \bar{x}_i = \sum_c x_{ic} = F(\sigma - 1)$$

The lets plug in the demands  $x_{ic}$  that we calculated above.

$$\begin{aligned} \bar{x} &= \sum_c x_{ic} = \sum_c p_i^{-\sigma} (T_{ic})^{1-\sigma} (\mu E_c) (P_c^M)^{\sigma-1} \\ &= \sum_c p_i^{-\sigma} (T_{ic})^{1-\sigma} (w_c L_c) (P_c^M)^{\sigma-1} \end{aligned}$$

which we can rewrite as

$$p_i^\sigma = \frac{1}{\bar{x}} \sum_c (T_{ic})^{1-\sigma} (w_c L_c) (P_c^M)^{\sigma-1}$$

or

$$\left( \frac{\sigma}{\sigma - 1} w_i \right)^\sigma = \frac{1}{\bar{x}} \sum_c (T_{ic})^{1-\sigma} (w_c L_c) (P_c^M)^{\sigma-1}$$

The authors refer to the equation above at the “tradeables wage equation.” Defining  $FMA_i$  (firm market access) as follows,

$$FMA_i \equiv \sum_c (T_{ic})^{1-\sigma} (w_c L_c) (P_c^M)^{\sigma-1}$$

we can solve for the wage

$$w_i = \xi [FMA_i]^{\frac{1}{\sigma}}$$

for a constant  $\xi$ .

### 4.3 Factor Market Equilibrium

Individuals will move across cities to that they are indifferent as to where they locate. The real wage is equalized,

$$\omega_c = \frac{w_c}{(P_c^M)^\mu (P_c^H)^{1-\mu}} = \omega$$

Labor market clearing implies

$$L_i = n_i \bar{l} = n_i F \sigma$$

(recall all firms in the economy have same volume from zero profit condition, so labor at each location is proportional to variety.

Let's substitute in all this interesting stuff we have been deriving.

$$\begin{aligned} \omega &= \frac{w_c}{(P_c^M)^\mu (P_c^H)^{1-\mu}} \\ &= \frac{w_c}{(P_c^M)^\mu \left( \frac{(1-\mu) w_c L_c}{H_c \mu} \right)^{1-\mu}} = \frac{w_c^\mu H_c^{1-\mu}}{(P_c^M)^\mu \left( \frac{(1-\mu)}{\mu} \right)^{1-\mu} L_c^{1-\mu}} \end{aligned}$$

Or

$$\begin{aligned} L_c &= H_c w_c^{\frac{\mu}{1-\mu}} (P_c^M)^{-\frac{\mu}{1-\mu}} \times \text{constant} \\ &= H_c [FMA_i]^{\frac{\mu}{\sigma(1-\mu)}} [CMA_c]^{\frac{\mu}{(1-\mu)(\sigma-1)}} \times \text{constant} \end{aligned}$$

### 4.4 General Equilibrium

- List of 7 variables for each city,  $w_c, p_c, L_c, n_c, P_c^M, P_c^H, E_c$
- Assume  $\sigma(1-\mu) > 1$ , they claim this is Helpman condition for unique equilibrium.

- Given any set of model parameters, can find  $H_c$  to perfectly fit the data. Fix parameters, there is a mapping between  $\{H_1, H_2, \dots, H_C\}$  and  $\{L_1, L_2, \dots, L_C\}$ . Fit 1939 prewar population.
- After the division, set  $T_{ic} = \infty$  for all cases that involve crossing the border. Solve for a new equilibrium distribution of population.

## 4.5 Parameterization

- $\sigma = 4$
- $\mu = \frac{2}{3}$  (so  $\sigma(1 - \mu) > 1$ )
- $T_{ic} = d_{ic}^\phi, \phi = \frac{1}{3}$
- Distance elasticity of trade  $(1 - \sigma)\phi = -1$ .
- only thing left is amenities...back them out.

## References

Costas Arkolakis, Arnaud Costinot, Andrés Rodríguez-Clare, “New Trade Models, Same Old Gains?” working paper, Sept 6, 2010.

Bernard, A. B., J. Eaton, J. B. Jensen and S. Kortum (2003), “Plants and productivity in international trade,” *American Economic Review*, 93(4): 1268-1290.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2009), “The productivity advantages of large cities: Distinguishing agglomeration from firm selection,” working paper, University of Toronto.

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

Helpman, E. (1998), “The size of regions,” in D.Pines, E.Sadka and Y.Zilcha (eds.), *Topics in Public Economics*, pp. 33-54, New York: Cambridge University Press.

Krugman, P. (1991), “Increasing Returns and Economic Geography”, *Journal of Political Economy* 99, 483-499.

Melitz, Marc J. (2003), “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 71(6): 1695-1725.

Melitz, Marc J. and G. Ottaviano (2008), “Market size, trade, and productivity,” *Review of Economic Studies*, 75(1): 295-316.

Ottaviano, G., T. Tabuchi and J.-F. Thisse (2002), “Agglomeration and trade revisited,” *International Economic Review*, 43(2): 409-435.

Stephen Redding and Daniel Sturm, “The Costs of Remoteness: Evidence from German Division and Reunification” *American Economic Review*, 98(5), 1766-1797, 2008