CESifo Working Papers
www.cesifo.org/wp

More Similar Firms – More Similar Regions? On the Role of Firm Heterogeneity for Agglomeration

Maximilian von Ehrlich
Tobias Seidel

CESIFO WORKING PAPER NO. 4028
CATEGORY 4: LABOUR MARKETS
DECEMBER 2012

CESifo *Working Paper No. 4028*

More Similar Firms – More Similar Regions? On the Role of Firm Heterogeneity for Agglomeration

Abstract

In contrast to what several papers have argued recently, we show that firm heterogeneity fosters agglomeration of economic activity. If firms are more similar with respect to their total factor productivity, each company faces a lower propensity to export. This renders the home market more important speaking against agglomeration. We also relate changes in firm heterogeneity to technological progress which allows us to derive novel insights on the role of technology for the location of economic activity.

JEL-Code: F120, F220, R120.

Keywords: firm heterogeneity, agglomeration, technological change, trade, labor mobility.

Maximilian von Ehrlich
ETH Zurich
Department of Management, Technology and Economics / WEH F9
Weinbergstrasse 35
8092 Zurich
Switzerland
vonehrlich@kof.ethz.ch

Tobias Seidel
ETH Zurich
Department of Management, Technology and Economics / WEH E4
Weinbergstrasse 35
8092 Zurich
Switzerland
seidel@kof.ethz.ch

November 2012

1 Introduction

Recent research in international economics has stressed the importance of heterogeneous firm productivities and their implications for trade flows, firm selection and welfare – to name just a few. These important insights in the trade context have stimulated researchers in economic geography to look more closely at the role of firm heterogeneity for agglomeration. In a recent survey article, Ottaviano (2011) highlighted that "further understanding how the heterogeneity of people and firms comes about and how it may help to shed light inside the black box of agglomeration economies is a very promising direction of future research."

We show in this paper that more heterogeneity of firms with regard to their total factor productivity is a driver of agglomeration. This is a novel insight as it stands in contrast to recent research in economic geography. By introducing heterogeneous firms in the tradition of Melitz (2003) into Krugman's (1991) core-periphery model, we account for exporter fixed costs and endogenous firm entry. This gives rise to firm selection effects that are crucial for our results. More dispersion in firm productivities raises the average efficiency of firms which drives the least productive companies out of the market and increases the probability of exporting for all surviving firms. We show that more heterogeneity in firm productivities reduces *all* agglomeration and dispersion forces. However, agglomeration forces become relatively stronger the more heterogeneous firms are. Hence, the answer to this paper's research question is an unequivocal "yes". If firms are more similar to each other in terms of productivity, the tendency for full agglomeration is weaker.

Our paper relates to several recent contributions in the economic geography literature that have taken up the notion of firm heterogeneity. For example, Baldwin and Okubo (2006) extend the footloose capital model to account for differences in firm productivity. They show that heterogeneity leads to a sorting of the most productive firms into larger regions. In contrast to our paper, however, firm heterogeneity does not affect the home-market effect and their model exhibits fewer economic channels than the richer core-periphery model that we use. Okubo, Picard and Thisse (2010) consider two types of firm productivity and show that the more productive firm type selects into large markets when trade costs fall, but more high-cost firms find it also profitable to locate there if trade costs fall even further. This gives rise to an inverted U-shape relationship between economic integration and the international productivity gap. Saito, Gopinath and Wu

(2011) employ a quadratic utility function and a fixed number of firms of two productivity types to arrive at the same conclusion with respect to sorting.[1] Further, Ottaviano (2011) concludes that heterogeneity works as a dispersion force in a duopoly framework with two regions because the less productive company "has a stronger incentive to avoid the tougher competition due to agglomeration in the advantaged location" (see also Behrens, Mion and Ottaviano, 2011).[2]

Our paper differs from this line of research in that we do not consider relocation of firms. Rather, and more in line with Krugman (1991), agglomeration works via exit and entry of firms in each region stimulated by labor mobility and thus changes in market size. We neither keep the number of firms fixed, but determine it endogenously in general equilibrium. Together with export selection effects, this characteristic is responsible for the opposite conclusions we derive. An important implication of this approach is that we do not have to make assumptions on which firms move first. The mobile factor (high-skilled labor) is homogeneous. As a result, spatial sorting does not occur in our model. Hence, our analysis compares more to recent work by Melitz and Ottaviano (2008) who show that less productive firms exit the larger market such that on average more productive firms produce there. A similar implication is contained in our model. Our paper is also related to Pflüger and Südekum (2013) who use a two-sector trade model with heterogeneous firms and asymmetric regions to examine the effects of subsidizing firm entry. However, the key difference is that they ignore international labor mobility and thus agglomeration.

A further important contribution of our paper is that we relate firm heterogeneity to technology based on the notion of first-order stochastic dominance. This concept was prominently introduced into the international trade literature by Demidova (2008) and allows us to derive novel insights on the link between technological progress and clustering of economic activity. We demonstrate that a lower shape parameter of the Pareto distribution implies higher average productivity draws such that technological progress (more heterogeneity) stimulates agglomeration. This is an important result since previous models based on homogeneous firms were unable to establish a link from technological progress to agglomeration.

[1] In an alternative model by Nocke (2006), the sorting of more productive firms in larger markets is driven by differences in managerial talent. However, his framework differs in that he models entry and exit decisions explicitly rendering the number of firms endogenous.

[2] Using a vertical-linkages model, Okubo (2009) finds that firm heterogeneity operates as an agglomeration force.

These insights contain important policy implications. As many politicians are concerned about structural policies to stimulate firm location in peripheral regions, it is key to understand how technological progress affects these location incentives. According to our model, productivity improvements that go in hand with a greater productivity dispersion foster agglomeration of industries rendering regional policies to achieve equality between jurisdictions more costly. Hence, the interdependence of productivity-enhancing policies and regional policies needs to be taken into account when designing the optimal policy mix.

We organize the paper in six parts. Section 2 lays out the model before we analyze the equilibrium in Section 3. Section 4 provides a discussion of the role of firm heterogeneity for the well-known agglomeration and dispersion forces. In Section 5, we introduce the notion of first-order stochastic dominance on which we base our discussion of technological change and agglomeration. Section 6 offers concluding remarks.

2 The model

Consider a world consisting of two regions, i and j, that are identical ex-ante. However, mobility of high-skilled workers H may give rise to concentration of economic activity in one region ex-post. This is referred to as a core-periphery agglomeration pattern as in Krugman (1991). Unless otherwise stated we report expressions for country i stressing that similar equations exist for region j.

2.1 Preferences and demand

Individuals derive utility from consuming two goods, a homogeneous good Y and a differentiated commodity X, based on

$$U_i = X_i^{\alpha} Y_i^{1-\alpha}. \tag{1}$$

The differentiated good is composed of an endogenously determined mass V of varieties v that are aggregated according to $X_i = \left(\int_{v \in V} \hat{x}_i\left(v\right)^{\frac{\sigma-1}{\sigma}} dv \right)^{\frac{\sigma}{\sigma-1}}$. The parameter $\sigma > 1$ represents the constant elasticity of substitution between any two varieties and $\hat{x}_i(v)$ denotes the consumption level of variant v. As there are two regions and the manufactured variety is generally tradable, $\hat{x}_i(v)$ may be a local or a foreign (imported) type.

Utility maximization yields total demand for variety v of the composite good,

$$\hat{x}_i(v) = \frac{p_i(v)^{-\sigma}}{P_i^{1-\sigma}} \alpha E_i, \tag{2}$$

where $P_i \equiv \left(\int_{v \in V} p_i(v)^{1-\sigma} dv\right)^{\frac{1}{1-\sigma}}$ denotes a price index of X_i, $p_i(v)$ represents the consumer price for variety v in region i and E_i is aggregate expenditure for consumption in region i. The expenditures for the differentiated and the homogeneous goods are $X_i = \alpha E_i / P_i$ and $Y_i = (1-\alpha) E_i / P_{Yi}$, respectively where P_{Yi} denotes the price of the homogeneous good in region i. Plugging the demand functions into (1) yields indirect utility

$$V_i = \frac{\alpha^\alpha (1-\alpha)^{1-\alpha} E_i}{P_i^\alpha P_{Yi}^{1-\alpha}}. \tag{3}$$

2.2 Technology and profits

Turning to the production side of the economy, we assume one unit of low-skilled labor L is required to produce one unit of the homogeneous good. Assuming further that trading this commodity is costless and choosing Y as the numéraire, we pin down low-skilled wages to unity in both regions. We disregard any asymmetries in the homogeneous good sector such that both regions are endowed with the same amount of low-skilled labor $L_i = L_j = L/2$.

In the manufacturing sector, firms demand high-skilled labor H as the only factor of production. Due to fixed costs and increasing returns to scale, each company produces a unique variety such that the market is characterized by monopolistic competition. We follow Melitz (2003) in assuming that firms differ in their labor productivity φ which is drawn from a commonly-known distribution function. Firms do not know their productivity ex ante, but have to incur an investment (like R&D) to obtain this information. We denote this entry cost in terms of high-skilled labor, that is $f_e w_i$. Based on this knowledge, firms decide about producing or exiting the industry in case their productivity is too low to make profits. As a result, each company has a specific input requirement per unit of output according to $x_i(\varphi) = \varphi h_i(\varphi)$, where $h_i(\varphi)$ denotes labor input by firm of type φ. Note that – as in Melitz (2003) – firms are heterogeneous with respect to their productivity although high-skilled workers do not differ in their skills. This can be rationalized by arguing that each firm possesses a specific technology determining labor productivity of each employee.

Furthermore, each company chooses which markets to serve. To be present on the local market, each company needs to invest f high-skilled workers as a fixed input. This investment could represent a distribution network or any marketing activity that is independent from the number of products sold. A similar argument applies for the export market such that firms have to hire f_x high-skilled workers to sell to customers abroad. While delivering goods to local customers is costless, exports imply transportation costs $\tau > 1$ per unit. Following the iceberg notation, this means that τ units have to be shipped for one unit to arrive at the final destination abroad.

Under Dixit-Stiglitz preferences, firms face a constant price elasticity of demand so that constant price-cost mark-ups ensure profit maximization. For domestic sales and exports (subindex x), consumer prices are respectively given by

$$p_i(\varphi) = \frac{\sigma w_i}{(\sigma - 1)\varphi} \qquad p_{ix}(\varphi) = \frac{\sigma \tau w_i}{(\sigma - 1)\varphi}.$$

Denoting by $H = H_i + H_j$ the aggregate endowment of high-skilled workers in both regions and by λ the corresponding share residing and working in region i, revenues and profits for each market are

$$r_i(\varphi) = \frac{p_i(\varphi)^{1-\sigma}}{P_i^{1-\sigma}} \alpha E_i \qquad r_{ix}(\varphi) = \frac{p_{ix}(\varphi)^{1-\sigma}}{P_j^{1-\sigma}} \alpha E_j$$

$$\pi_i(\varphi) = r_i(\varphi)/\sigma - f w_i \qquad \pi_{ix}(\varphi) = r_{ix}(\varphi)/\sigma - f_x w_i,$$

where $E_i = w_i \lambda H + L/2$ and $E_j = w_j(1 - \lambda)H + L/2$. Notice that firms with higher productivity charge lower prices, sell more and earn higher profits.

We follow the literature on heterogeneous firms in assuming Pareto-distributed productivity levels. The cumulative distribution function reads $G(\varphi) = 1 - \varphi^{-k}$, where $k > 0$ denotes the shape parameter. Note that we have normalized the scale parameter to unity without loss of generality to simplify notation. This means that $\varphi = 1$ is the lowest productivity a firm can draw. The Pareto distribution offers the advantage that the shape parameter k is a straightforward measure for the similarity of firms. The variance of the Pareto distribution $Var(\varphi) = k/[(k-1)^2(k-2)]$ is strictly decreasing in k for $k > 2$.[3] A

[3] Assuming $k > 2$ is necessary for ensuring that the Pareto distribution has a finite variance. See also Helpman, Melitz, and Yeaple (2004). Further, we impose $k > \sigma - 1$ to ensure that the integrals for the

high value of k implies that it becomes less likely to draw a high productivity level φ. In other words, there are only a few very productive firms and many low-productive ones. In the boundary case of $k = \infty$, all firms are clustered at the lower bound $\varphi = 1$. We thus refer to higher levels of the shape parameter as a more similar distribution of productivity levels.

3 Equilibrium

The timing of the model is as follows: For a given allocation of labor, firms decide to enter the industry until expected profits are zero. Based on their productivity draw, they start producing as long as their profits are not negative and pay the market wage w. This is true for all firms with a productivity level φ that exceeds the cutoff level φ^*. A subset of these domestically active firms will find it profitable to export to the foreign region. We refer to this situation as the short-run equilibrium without labor mobility. In the long run, high-skilled workers choose their residence in either region based on real wages. If migrating to the other region promises higher real remuneration, the allocation of workers is modified such that firm entry and exit adjusts to meet the equilibrium condition of zero expected profits. The migration process terminates if either real wages are equated across regions (symmetric equilibrium) or all workers cluster in one jurisdiction (core-periphery equilibrium). In contrast to most of the previous contributions in this literature, our model features an endogenous number of firms as well as an endogenous average productivity.

In the following, we first determine the short-run equilibrium. To do this, it is essential to derive the domestic cutoff productivity level φ_i^*. Then, the only endogenous variables left are the number of firms and the wage rates. To obtain φ_i^*, we combine the free-entry condition (FE) with the zero-cutoff-profit condition (ZCP). Firms enter the industry as long as expected profits (from both domestic sales and exports) are sufficiently high to cover the fixed market entry costs. Formally, this condition is given by

$$\left(\varphi_i^*\right)^{-k} \bar{\pi}_i = f_e w_i, \qquad (FE)$$

where $\bar{\pi}_i$ denotes average profits of surviving firms. If this term is multiplied by the probability of surviving in competition, that is $1 - G\left(\varphi_i^*\right) = \left(\varphi_i^*\right)^{-k}$, we obtain expected

average productivity of the Pareto distribution converge.

profits before firm-specific productivity levels have been realized.

Referring to $\tilde{\varphi}_i$ and $\tilde{\varphi}_{ix}$ as the productivity levels of the average domestic and exporting firm, respectively, surviving firms can expect to earn $\pi_i(\tilde{\varphi}_i)$ domestically and $(\varphi_i^*/\varphi_{ix}^*)^k \pi_{ix}(\tilde{\varphi}_{ix})$ from exports. The term $(\varphi_i^*/\varphi_{ix}^*)^k$ reflects the probability of becoming an exporter conditional on being active in the domestic market, with φ_{ix}^* denoting the export productivity cutoff. Firms will only start producing for the domestic and the export market as long as their revenues from the respective market cover the market-specific fixed costs. Hence, the marginal domestic and exporting firm will be characterized by

$$r_i(\varphi_i^*) = \sigma f w_i \qquad \text{and} \qquad r_{ix}(\varphi_{ix}^*) = \sigma f_x w_i.$$

These two conditions can be used together with $r_{jx}(\varphi_{jx}^*) = r_i[(\varphi_{jx}^* w_i)/(\tau w_j)] = \sigma f_x w_j$ to establish a link between the domestic cutoff in region i and the exporter cutoff in region j:

$$\varphi_{jx}^* = \tau (f_x/f)^{1/(\sigma-1)} (w_j/w_i)^{\sigma/(\sigma-1)} \varphi_i^*. \tag{4}$$

We restrict our analysis to the realistic scenario where domestic sales are generally more profitable than exporting which requires $f_x > f$. This is a common assumption in the literature to avoid the case where firms export without serving local consumers.[4] For asymmetric regions, ensuring that firms also serve the domestic market when exporting – i.e. $\varphi_{jx}^* > \varphi_j^*$ and $\varphi_{ix}^* > \varphi_i^*$ – imposes a limit on relative wages in the two regions.[5] These constraints on relative wages are similar to the constraint Demidova (2008) imposes on asymmetric productivity distributions. With these insights at hand, it is evident that the conditional export probability is limited to range between zero and unity, with higher levels of the shape parameter k – i.e. more similar productivity distributions – implying a lower export probability. Using (4), we can formulate the conditional export probability as

$$\left(\frac{\varphi_i^*}{\varphi_{ix}^*}\right)^k = \tau^{-k} \left(\frac{f}{f_x}\right)^{\frac{k}{\sigma-1}} \left(\frac{w_j}{w_i}\right)^{\frac{\sigma k}{\sigma-1}} \left(\frac{\varphi_i^*}{\varphi_j^*}\right)^k. \tag{5}$$

It is immediate that the probability of exporting is decreasing in f_x and τ and increasing in f and k. Further, a higher relative wage in region j reduces that region's competitiveness

[4]Note that for symmetric regions $w_i = w_j$ and $\varphi_i^* = \varphi_j^*$ such that $\varphi_{jx}^* > \varphi_i^*$ implies $\varphi_{jx}^* > \varphi_j^*$ and $\varphi_{ix}^* > \varphi_j^*$ implies $\varphi_{ix}^* > \varphi_i^*$.

[5]The necessary and sufficient conditions for ensuring that only domestically active firms start exporting are derived in a supplement to this paper.

such that exporting for firms based in region i becomes more likely (for a given productivity φ). We will rely on this equation later as the export probability plays a central role in understanding the economic channels underlying our model. Assuming a Pareto distribution implies that average productivities result as a constant markup over the respective cutoff levels, that is $\tilde{\varphi}_i/\varphi_i^* = \tilde{\varphi}_{ix}/\varphi_{ix}^* = [k/(k-\sigma+1)]^{1/(\sigma-1)}$ which enables us to state average profits in terms of the cutoff productivities. Combining the profits from domestic and export sales with the conditional export probability in (4), we can express the zero-cutoff-profit condition as[6]

$$\bar{\pi}_i = \frac{\sigma-1}{k-\sigma+1}\left[1+\phi\left(\frac{w_j}{w_i}\right)^{\frac{\sigma k}{\sigma-1}}\left(\frac{\varphi_i^*}{\varphi_j^*}\right)^k\right] f w_i, \qquad (ZCP)$$

where $\phi \equiv \tau^{-k}\left(f/f_x\right)^{\frac{k-\sigma+1}{\sigma-1}}$ can be interpreted as a measure of trade freeness (see Pflüger and Südekum, 2013). Relative wages play a role for average profits as a higher wage level abroad renders that region less competitive and domestic exporters can penetrate the foreign market more easily.

Using (FE) and (ZCP) for each region delivers two equations that can be solved for the domestic cutoff in i,[7]

$$\varphi_i^* = \left[\frac{\sigma-1}{k-\sigma+1}\frac{1-\phi^2}{1-\phi\left(\frac{w_j}{w_i}\right)^{\frac{\sigma k}{\sigma-1}}}\frac{f}{f_e}\right]^{\frac{1}{k}}. \qquad (6)$$

The region with the higher wage rate features the lower cutoff productivity because higher wages reduce expected profits and result in less entry.

We finally need to solve for the number of firms M_i and M_j, as well as for the nominal wage rates, w_i and w_j. To do so, we employ the respective labor-market-clearing conditions jointly with the trade-balance equations. Expressing the number of exporters and the number of entrants respectively as $M_{ix} = \left(\varphi_i^*/\varphi_{ix}^*\right)^k M_i$ and $M_{ie} = \left(\varphi_i^*\right)^k M_i$, we can

[6]See Appendix A for further details on the derivation of this condition.
[7]Appendix B provides further details on the derivation of this equation.

formulate the market-clearing condition for high-skilled workers in region i as

$$\lambda H = M_i \left[\frac{\hat{x}_i (\tilde{\varphi}_i)}{\tilde{\varphi}_i} + f \right] + \left(\frac{\varphi_i^*}{\varphi_{ix}^*} \right)^k M_i \left[\frac{\tau \hat{x}_{ix} (\tilde{\varphi}_{ix})}{\tilde{\varphi}_{ix}} + f_x \right] + (\varphi_i^*)^k M_i f_e. \tag{7}$$

The trade-balance condition equates net exports of manufactured varieties with net imports of the homogeneous good. The latter is the difference between agricultural production, L_i, and expenditure, $(1-\alpha)[L/2 + \lambda H w_i] - L_i$, and we have

$$(1-\alpha) \lambda H w_i - \alpha L_i = \left(\frac{\varphi_i^*}{\varphi_{ix}^*} \right)^k M_i r_{ix} (\tilde{\varphi}_{ix}) - \left(\frac{\varphi_j^*}{\varphi_{jx}^*} \right)^k M_j r_{jx} (\tilde{\varphi}_{jx}). \tag{8}$$

Note that consumer demand and revenues are functions of the price index which is given by

$$P_i = \left[M_i \left(\frac{\sigma w_i}{(\sigma-1) \widetilde{\varphi}_i} \right)^{1-\sigma} + \left(\varphi_j^* / \varphi_{jx}^* \right)^k M_j \left(\frac{\tau \sigma w_j}{(\sigma-1) \widetilde{\varphi}_{jx}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{9}$$

Similar expressions exist for region j.

As in Krugman's core-periphery model, endogenous variables enter in a non-linear fashion such that closed-form solutions for the full range of allocations of high-skilled workers are generally infeasible. Yet, we can solve the model analytically for the three specific allocations of high-skilled workers which turn out as equilibrium candidates in the long-run analysis below that is for $\lambda = 1/2$ and $\lambda = 1 \vee \lambda = 0$.[8] This enables us to derive closed-form solutions for the critical level of trade costs where the symmetric equilibrium becomes unstable and where agglomeration starts to represent a stable outcome. As in the literature, we refer to these values as the break point τ_B and the sustain point τ_S, respectively. In order to solve the model for all possible values of λ, we choose parameters and focus on variations in trade costs and firm heterogeneity as measured by k.[9]

[8] See Appendix C for the equilibrium values of wages and firm numbers in the $\lambda = 1/2$ and $\lambda = 1 \vee \lambda = 0$ scenarios.

[9] To solve the model, we follow the standard procedure in first determining the short-run equilibrium (without labor mobility) to obtain values for nominal wages and the number of firms. This enables us in a second step to compute the difference in indirect utilities to determine the migration and hence, the agglomeration pattern in the long run. In choosing parameter values, we account for the 'no-black-hole'-condition, that is $(\sigma-1)/\sigma > \alpha$.

4 How firm heterogeneity affects agglomeration

We are now ready to examine the role of firm heterogeneity for long-run industry agglomeration equilibria. To illustrate such equilibria and how firm heterogeneity affects them, we compute the utility difference of a mobile high-skilled worker across both regions for given λ,

$$\Delta \equiv V_i - V_j = \frac{w_i}{P_i^\alpha} - \frac{w_j}{P_j^\alpha}. \tag{10}$$

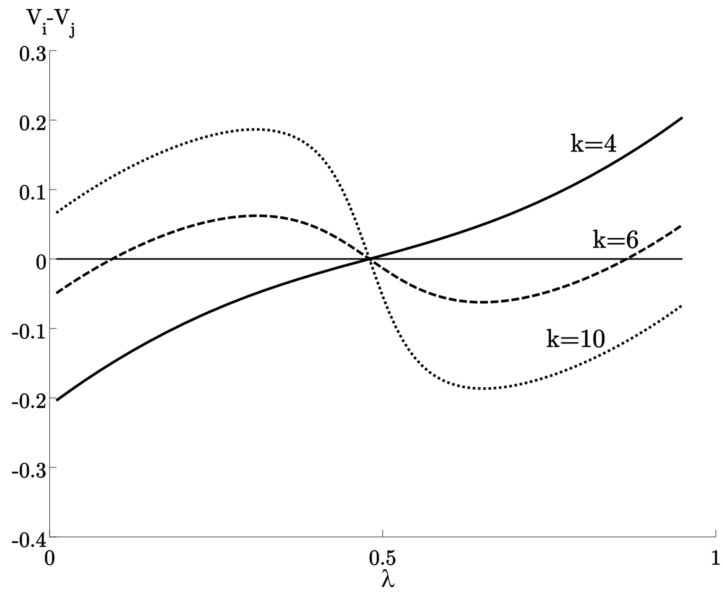
It is immediate from (10) that a combination of nominal wages and consumer prices determines this utility difference. In the traditional new economic geography literature, Δ crucially depends on the level of trade costs between regions i and j. High levels of trade costs usually imply a stable symmetric interior equilibrium ($\partial\Delta/\partial\lambda|_{\lambda=0.5} < 0$) while industry concentrates in one region entirely at low trade barriers ($\partial\Delta/\partial\lambda|_{\lambda=0.5} > 0$). Although we account for selection effects due to firm heterogeneity, our model establishes the same pattern as in the standard core-periphery model (see also von Ehrlich and Seidel, 2011).[10] We can thus focus on the role of productivity improvements and choose *one* level of τ and *various* distribution shape parameters k with higher levels capturing more similarity of firms.

Figure 1 plots equilibria for three alternative values of k and each potential (fixed) distribution of high-skilled workers λ in both regions.[11] This illustration helps identify stable equilibria if high-skilled workers choose their region of residence. An interior equilibrium requires the utility difference defined in (10) to equal zero. A symmetric distribution of H represents always an equilibrium, but it is not generally stable. Stability requires that $\partial\Delta/\partial\lambda|_{\lambda=0.5} < 0$ which is satisfied for $k = 10$, for example. Moving one high-skilled worker from region j to region i lowers indirect utility in i causing re-migration to j. Raising firm heterogeneity by choosing a lower shape parameter $k = 6$ rotates the Δ-function in Figure 1 anti-clockwise around $\lambda = 0.5$. While the symmetric interior equilibrium is still stable, there are two additional asymmetric λ-values that satisfy $V_i = V_j$, but they are unstable. A further reduction of k rotates the Δ-function further implying a core-

[10] These results are available from the authors upon request.

[11] We have chosen the following parameters for this simulation: $H = 100$, $L_i = L_j = 150$, $\tau = 1.5$, $\alpha = 0.4$, $\sigma = 3.8$, $f = 1$, $f_x = 3$ and $f_e = 0.2$. The effects of heterogeneity on the agglomeration pattern remain unaffected by the choice of these parameters. Results for alternative parameter constellations can be provided by the authors upon request.

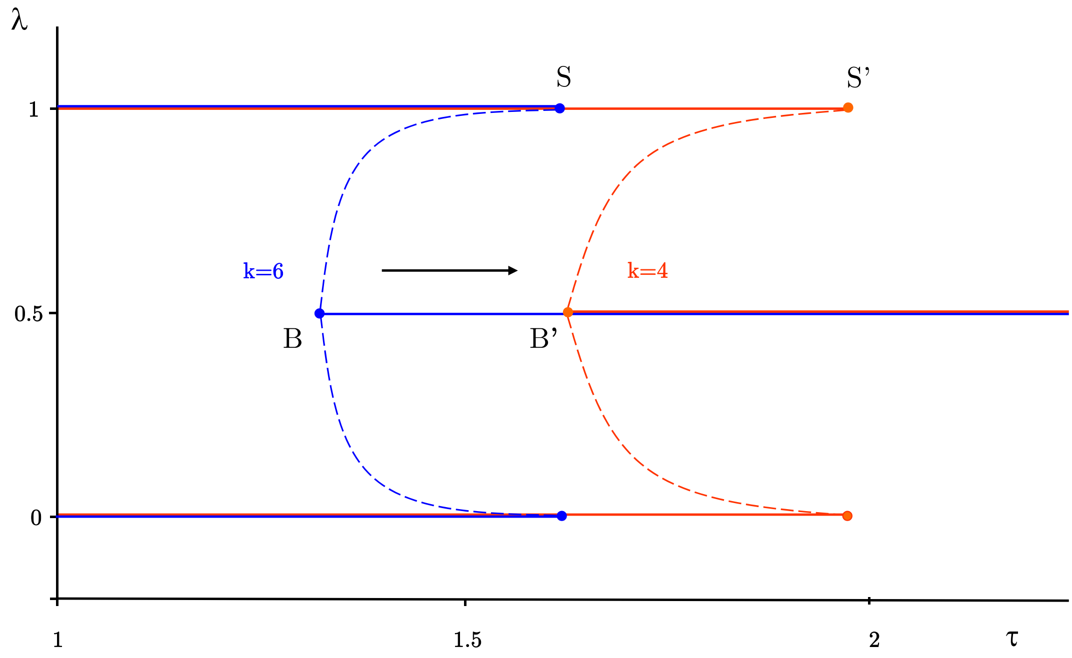
Figure 1: Equilibria and firm heterogeneity



periphery-agglomeration pattern for $k = 4$. Migration from j to i now raises indirect utility in region i stimulating further migration until all high-skilled workers and thus all manufacturing firms locate in that region. However, it is indeterminate which region becomes the core due to symmetry. In sum, raising firm heterogeneity fosters agglomeration.

The informed reader has realized that Figure 1 looks identical to the case where trade costs vary. Before we lay out the intuition why more heterogeneity between firm productivities has similar effects as lower trade costs, we first want to derive the stable long-run agglomeration equilibria according to the conditions explained above. In contrast to Figure 1 where we have analyzed the effects of productivity improvements for a certain level of trade costs, we now focus on how a change in k affects the allocation of economic activity during a process of falling trade costs. Figure 2 summarizes two scenarios with different degrees of firm heterogeneity, $k = 6$ (blue) and $k = 4$ (orange). The solid lines indicate stable equilibria while the dashed lines represent asymmetric, but unstable interior equilibria. There are two critical levels of trade costs: the so-called *break* and *sustain points* indicated in Figure 2 by B and by S, respectively. The *break point* refers to the level of trade costs below which the symmetric equilibrium is unstable while the *sustain point*

Figure 2: Industry location, trade costs, and heterogeneity



refers to the level of trade costs below which a full-agglomeration equilibrium is stable. We observe from Figure 2 that full agglomeration occurs as a stable equilibrium at higher levels of trade costs when firms in both regions are less similar (i.e. when k is lower). At the same time, symmetry ceases to be a stable equilibrium at higher levels of trade costs when the degree of heterogeneity is greater. More heterogeneity shifts the break point towards higher levels of trade costs.

We can also show this result analytically. The break point corresponds to the level of trade costs which fulfills

$$\frac{d\Delta}{d\lambda}|_{\lambda=1/2} = 0. \tag{11}$$

Using the total differentials of the price index (9), the labor market constraint (7), and the trade balance condition (8) in combination with the equilibrium values of wages and firm numbers at $\lambda = 1/2$, we get[12]

$$\tau_B = \left(\frac{f}{f_x}\right)^{\frac{k-\sigma+1}{k(\sigma-1)}} \Theta^{-\frac{1}{k}}, \tag{12}$$

where

$$\Theta = \frac{1 + 2k\alpha^2\sigma^2 + \alpha^2\sigma - 2\sigma - \left(\alpha^2 - 1\right)\sigma^2 - \alpha\sqrt{1 - 2\sigma + \sigma^2 + 4k\sigma + 4k\sigma^2\left(\alpha^2 + \alpha^2 k - \alpha^2\sigma + \sigma - 2\right)}}{(1+\alpha)(\sigma-1)(\sigma - 1 + \alpha\sigma)}$$

ranges between zero and one.[13] Taking the derivative of (12) with respect to k reveals that higher levels of k reduce the break point τ_B. We prove this result formally in Appendix C where we also provide an implicit solution for the sustain point τ_S.

To understand the role of firm heterogeneity for the incentive to agglomerate in one region, we briefly discuss the underlying agglomeration and dispersion forces in Krugman's agglomeration model. This serves as a starting point to study how firm heterogeneity affects each of these channels. First, mobile workers prefer to reside in the region associated with the lower consumer price index. At given nominal wages, this increases their real income and thus utility. As the larger region is characterized by lower prices (due to fewer imports), the *price-index effect* works in favor of agglomeration. Further, nominal wages are a crucial determinant for the migration decision. They are affected by two forces, the *home-market effect* and the *competition effect.* The former implies that firms benefit from producing in the larger market as they earn higher profits from domestic sales in this region while export profits are lower (as transport costs reduce demand) and less important as the other region is relatively smaller. Hence, the larger market hosts a more than proportional number of firms (Krugman, 1980). This channel leads to higher nominal remuneration to labor as firms bid up wages more due to higher profits. Higher wages stimulate migration to the larger region and thus work in favor of agglomeration. On the other hand, firms face more competition and each firm gets a lower market share. This implies lower profits such that firms would have to cut on wage costs. As the competition effect leads to lower nominal wages, it works as a stabilizing dispersion force in the core-periphery model.

It is well understood that a reduction in trade costs affects these forces to different

[12] Derivation details are provided in a supplement to this paper.

[13] Note that $\Theta = \phi(\tau_B)$ and therefore is limited to range between zero and unity. It is evident from (12) that symmetry could in principle be stable for all levels of trade costs $\tau > 1$. This would imply that the break point is less than unity. Such a scenario occurs if the export fixed costs, f_x, are very high relative to the fixed costs for domestic production, f. We thus assume that $f/f_x > \Theta^{\frac{\sigma-1}{k-\sigma+1}}$.

extents such that full agglomeration becomes a stable equilibrium at low levels of τ. In our model, lower trade costs do not only work via a price effect in that consumer prices for imported varieties decline, but also through an exporter selection effect. Reducing trade barriers also allows less productive firms to earn positive profits from serving customers abroad. This implies a higher propensity to export and raises the share of available varieties in each market. Further, intensified competition from high-productive companies drives the least productive firms out of the market and thereby raises average output per firm.

Firm heterogeneity works in a very similar way as the selection effects that come along with trade liberalization. Lowering k makes it more likely for a firm to draw a high productivity level which leads to more efficient competitors. Accordingly, the least productive firms are forced to exit and a higher share of companies that survive finds it profitable to pay the fixed costs for exporting. This increase in the probability to export impacts the balance between agglomeration and dispersion forces. We have shown that lower levels of k strengthen the two agglomeration forces relative to the dispersion force such that the critical trade cost level τ_B increases.

Let us now study the impact of firm heterogeneity on each of these three channels in more detail. As productivity improvements increase the propensity to export, it becomes less important to live in the larger region because a larger fraction of globally produced varieties is now available in each market. Hence, a decrease in k attenuates the price-index effect. The home-market effect as the second agglomeration force is also weakened with more heterogeneous firms because firms are shielded less from foreign competitors (export propensity is higher). As a consequence, the share of profits earned from domestic sales is lower rendering this market less important for overall profits. This further implies that firms have to reduce their wage payments as their location advantage has deteriorated. Finally, as a decrease in k leads to fewer but more efficient firms, each company owns a larger market share such that the migration of one worker from market i to market j reduces the market shares of all incumbent firms less. As we know from our comparative static results that the break point is decreasing in k, we conclude that a decrease in firm heterogeneity weakens the agglomeration forces less than the dispersion force. In sum, more heterogeneity among firms operates as a driving force for the clustering of industrial activity.

5 A notion of technological progress

Rather than interpreting the shape parameter as a measure of heterogeneity only, we can alternatively relate k to exogenous changes in technology. Recall that the average productivity under the Pareto assumption is given by a markup $[k/(k-\sigma+1)]^{1/(\sigma-1)}$ over the respective productivity cutoff. Importantly, this markup is decreasing in k. Hence, an increase in heterogeneity is associated with higher average productivity. In other words, a lower shape parameter implies a higher probability for every firm to draw a better productivity parameter φ. In the following, we apply the concept of hazard rate stochastic dominance (HRSD) which has been used in the international trade literature (see, for example, Demidova, 2008). We apply this concept to describe improvements in productivity by a reduction in k.

Definition A productivity distribution $G_1(\varphi)$ dominates the distribution function $G_0(\varphi)$ according to the hazard rate order $G_1(.) >_{hr} G_0(.)$ if for any productivity level φ,

$$\frac{g_1(\varphi)}{1-G_1(\varphi)} < \frac{g_0(\varphi)}{1-G_0(\varphi)}.$$

The hazard rate reflects the probability of observing a productivity φ, conditional on the productivity being above φ. Hence, if $G_1(.)$ dominates $G_0(.)$ according to the hazard rate order, this implies that firms facing a distribution $G_1(.)$ have a strictly higher probability of drawing a productivity above a certain level φ than firms facing $G_0(.)$.[14] We know from the discussion above that profits strictly increase in productivity φ, so any potential entrant would prefer to draw its productivity level from a distribution function that delivers a realization of at least φ with a higher probability. Under the Pareto specification, it is straightforward to see that ceteris paribus the distribution function with the lower shape parameter k dominates in terms of hazard rate order. Accordingly, we define technological progress as a reduction in k. Falvey, Greenaway and Yu (2006, 2011) consider the weaker concept of first order stochastic dominance which is not sufficient for the channel we highlight in our model. If $G_1(.)$ dominates $G_0(.)$ in the first-order sense this does not necessarily imply that the export probability conditional on surviving the entry

[14] Note that HRSD implies first order stochastic dominance (FSD) such that for any function $f(y)$ which is increasing in y, $G_1(.) >_{hr} G_0(.)$ implies $E_1[f(y)] > E_0[f(y)]$. Being the stricter concept than FSD, HRSD implies in addition that the expectations conditional on y exceeding a certain threshold φ dominate, i.e. $E_1[f(y)|y>\varphi] > E_0[f(y)|y>\varphi]$.

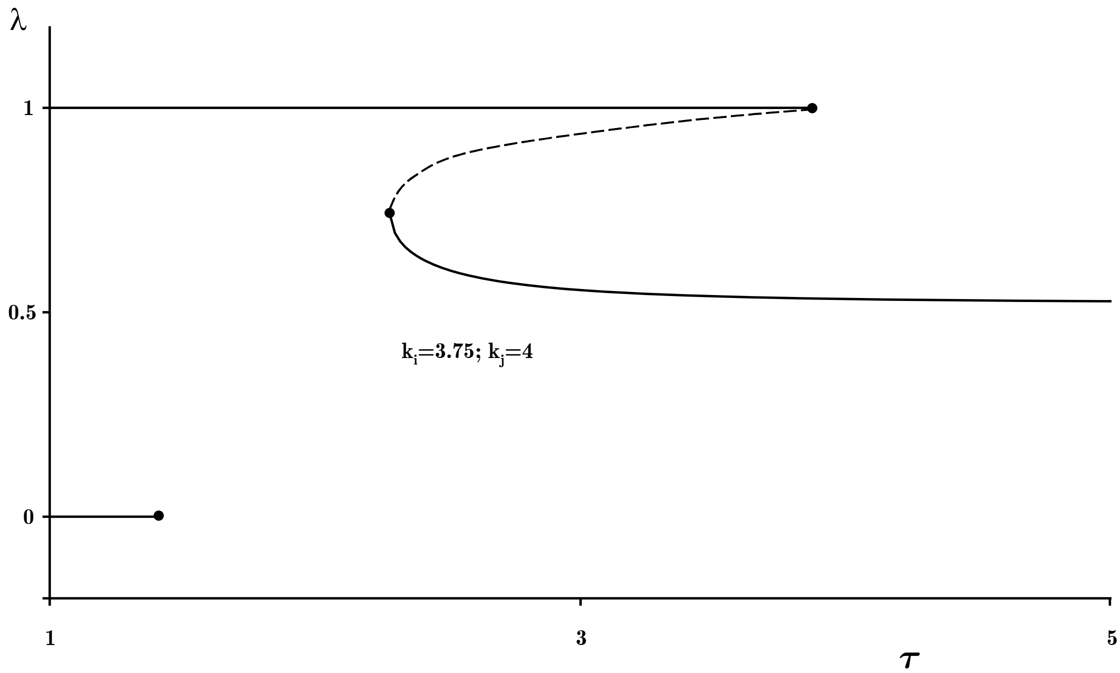
lottery increases which is key to our argument. According to our definition, technological progress not only increases the chances of being able to pay the fixed costs, but also raises the expected productivity conditional on being able to enter the market. Even though the expected productivity increases, the lowest possible productivity draw remains constant such that technological progress is reflected by a different shape of the distribution rather than by a different location. Empirical studies by Dunne, Foster, Haltiwanger, and Troske (2004) and by Faggio, Salvanes, and van Reenen (2010) provide evidence that technological progress is typically accompanied by an increase in the productivity dispersion. A theoretical argument for the positive correlation between average productivity and productivity dispersion could be made for example by assuming differences in the firm-specific adoption rates with regard to the use of new machines and technologies (see Caselli, 1999).

Relating firm heterogeneity to technology allows us to derive novel insights. Exogenous technological progress has been largely ignored in the new economic geography literature which is likely due to the fact that a symmetric increase in each region's productivity exerts no effect on the long-run location pattern in traditional agglomeration models with homogeneous firms. Introducing firm heterogeneity into the core-periphery model alters this implication and provides a new perspective on the role of technological progress for agglomeration. Due to firm heterogeneity and the selection effects discussed above we can identify a role for technology: Technological progress fosters agglomeration.

Technology is important because it is regarded as a key driver behind economic growth. This warrants a center-stage role in economic analysis and makes it a primary objective for public policies. On the one hand, the innovation and adoption of new technologies is likely to be facilitated by a higher density of employment and production as this gives rise to geographically localized externalities. On the other hand, technological progress itself may shape geography as higher levels of productivity raise the share of exporting firms and strengthen their tendency to cluster in space. While the former relationship has been emphasized in a strand of papers connecting new economic geography to endogenous growth theory (see Martin and Ottaviano, 1999, and Baldwin and Forslid, 2000), the latter relationship has not been topical in the literature.

Based on the technology interpretation of the shape parameter, it is interesting to also consider asymmetric technological progress. Assuming that region i experiences technological progress (that is a reduction in k_i while k_j remains constant), we observe that industry will now unambiguously cluster in the technologically advanced region in the pro-

Figure 3: Trade costs and asymmetric technologies



cess of falling trade costs, as shown in Figure 3.[15] With more productive firms in region i, the utility differential $V_i - V_j$ shifts upwards for all levels of trade costs (see Figure 4 in Appendix D).

This implies that symmetric interior equilibria at $\lambda = 0.5$ no longer exist. Instead, stable interior equilibria emerge with $\lambda > 0.5$. Moreover, the industry share monotonically increases as trade costs decline up to the point where a marginal reduction in τ makes all remaining high-skilled workers in region j move to region i (referred to as the break point).[16] Note that in such a situation full agglomeration in j can only occur for very low levels of trade costs. However, during an integration process this outcome is rather implausible since full agglomeration in i will become the only stable equilibrium once trade

[15]We have chosen the same parameters for this simulation as in the symmetric case. In a footloose entrepreneur model with imperfect labor markets and unemployment, Egger and Seidel (2008) derive a similar pattern of agglomeration when one region has more severe labor market frictions than the other.

[16]The dashed line indicates unstable interior equilibria that may occur at medium levels of trade costs. The $\tau = 2.4$ panel in Figure 4, Appendix D, illustrates these unstable interior equilibria.

costs fall below the break point. A further reduction will render agglomeration in j at some point stable as well, but it would require a sudden shift of the majority of workers from i to j in order to reach this equilibrium. Accordingly, the location of the core becomes determined once slight regional productivity differences exist.

These insights have important policy implications as we can derive a conflict of goals from our model. On the one hand, politicians are concerned about regional disparities and favor redistribution of income form rich (agglomerated) regions to poor (peripheral) regions. On the other hand, technological progress ranks very high on the agenda calling for policies that stimulate R&D to boost productivity. However, we have seen that stimulating higher (average) productivity corresponds to a more unequal distribution of firm productivities which in turn fosters agglomeration forces and makes regions more dissimilar. A way to circumvent this conflict is to focus on industries characterized by less advanced technologies (higher k) when aiming at stimulating technological progress in the periphery. In this regard, our model provides a rationale for industry-specific policies.

6 Concluding remarks

In this paper, we show that firm heterogeneity is a driver for the agglomeration of economic activity. This effect is primarily caused by the selection of firms into exporters and non-exporters and the endogenous adjustment of the number of firms. If firms differ more in terms of their total factor productivity, smaller and less efficient firms will be driven out of the market rendering average profits from domestic sales and exports higher for all surviving firms. Interestingly, trade liberalization exerts a similar effect on the propensity to export and it is well known that full agglomeration becomes a stable equilibrium for sufficiently low barriers to trade.

Our results contrast with recent research in economic geography that highlights the sorting of firms into different markets, but identifies no role of firm heterogeneity for relative market size. This is a key contribution of our paper, however. With more similar firms there is less incentive for high-skilled workers and firms to cluster in one region. The economic mechanism works as follows: As more heterogeneity among firms raises the probability of becoming an exporter, the home-market effect is attenuated because export profits become relatively more important. Similarly, the price index differs less across regions because a higher share of globally produced varieties becomes available in each region. Similarly to trade liberalization, we have shown that heterogeneity reduces the local competition effect more than the two agglomeration forces such that concentration of economic activity in one region becomes more likely.

We further link heterogeneity of firms to technology by interpreting the shape parameter of the productivity distribution as a measure for the level of technology. We can do this as a lower shape parameter is associated with higher average productivity. This allows us to make novel statements about the role of technology for agglomeration. This was generally infeasible in models with homogeneous firms because symmetric reductions in unit input coefficients turned out to neutralize each other across regions. In the presence of heterogeneous firms, however, selection effects render symmetric changes in the technology parameter no longer neutral. Instead, identical advances in technology in *both* regions make agglomeration a more likely outcome.

This result has important policy implications. As politicians are usually concerned about both a policy that favors technological progress and a policy that supports location of industry in peripheral regions, our results suggest that redistribution between regions becomes a more costly policy in the presence of stronger agglomeration forces. It is thus crucial to account for this interdependence in choosing a balance between both policy goals.

Appendix

A Deriving the zero-cutoff-profit (ZCP) condition

Average expected profits accrue from domestic operating profits $r_i\left(\widetilde{\varphi}_i\right)/\sigma$ minus fixed costs for domestic sales and from foreign operating profits $r_{ix}\left(\widetilde{\varphi}_{ix}\right)/\sigma$ minus export fixed costs. The latter have to be weighted by the probability of becoming an exporter conditional on being active in the domestic market. Hence, the average expected profits in region i are given by

$$\overline{\pi}_i = \frac{r_i\left(\widetilde{\varphi}_i\right)}{\sigma} - f w_i + \left(\varphi_i^*/\varphi_{ix}^*\right)^k \left[\frac{r_{ix}\left(\widetilde{\varphi}_{ix}\right)}{\sigma} - f_x w_i\right].$$

Substituting $r(\widetilde{\varphi}_i) = \left(\frac{\widetilde{\varphi}_i}{\varphi_i^*}\right)^{\sigma-1} r(\varphi_i^*)$ and $r(\widetilde{\varphi}_{ix}) = \left(\frac{\widetilde{\varphi}_{ix}}{\varphi_{ix}^*}\right)^{\sigma-1} r(\varphi_{ix}^*)$ as well as the zero-profit conditions $r(\varphi_i^*) = \sigma f w_i$ and $r(\varphi_{ix}^*) = \sigma f_x w_i$ yields

$$\overline{\pi}_i = \left[\left(\frac{\widetilde{\varphi}_i}{\varphi_i^*}\right)^{\sigma-1} - 1\right] f w_i + \left(\frac{\varphi_i^*}{\varphi_{ix}^*}\right)^k \left[\left(\frac{\widetilde{\varphi}_{ix}}{\varphi_{ix}^*}\right)^{\sigma-1} - 1\right] f_x w_i.$$

Finally, using the fact that under Pareto-distributed productivities $\widetilde{\varphi}_i/\varphi_i^* = \widetilde{\varphi}_{ix}/\varphi_{ix}^* = [k/(k-\sigma+1)]^{1/(\sigma-1)}$ together with (5) delivers (ZCP).

B Domestic cutoffs

Using the zero-cutoff-profit conditions (ZCP) together with the free entry conditions (FE) yields

$$\frac{(k-\sigma+1) f_e}{(\sigma-1) f} = \left[(\varphi_i^*)^{-k} + \phi \left(\frac{w_j}{w_i}\right)^{\frac{\sigma k}{\sigma-1}} (\varphi_j^*)^{-k}\right]$$
$$\frac{(k-\sigma+1) f_e}{(\sigma-1) f} = \left[(\varphi_j^*)^{-k} + \phi \left(\frac{w_i}{w_j}\right)^{\frac{\sigma k}{\sigma-1}} (\varphi_i^*)^{-k}\right]$$

for regions i and j, respectively. This equation system can be solved straightforwardly for the domestic cutoffs in each region.

C Break point, sustain point and firm heterogeneity

To derive the break and sustain point we need the equilibrium values of wages and firm numbers in the symmetric equilibrium as well as in the agglomeration equilibria. Without loss of generality we normalize in the following the total number of high skilled workers to $H = \alpha$ and the total number of low-skilled workers to $L = (1-\alpha)$.[17] Using (8) jointly with (7) and the cutoff productivities from (6), we can show that the symmetric equilibrium with $\lambda = 1/2$ is characterized by

$$w_i = w_j = 1 \qquad M_i = M_j = \frac{\alpha(k-\sigma+1)}{2f(1+\phi)k\sigma}. \tag{13}$$

[17] See Baldwin, Forslid, Martin, Ottaviano, and Robert-Nicoud (2003) on a discussion of normalizations in the core-periphery model. The analysis could be conducted without normalizing H/L but the analytics simplify a lot due to the normalizations. We refrain from this normalization in the numerical simulations.

Similarly, we can derive the equilibrium wages and firm numbers for an allocation with all high-skilled laborers residing in region i, i.e. for $\lambda = 1$:[18]

$$
\begin{aligned}
w_i &= 1, \qquad\qquad M_i = \frac{\alpha(1+\alpha)(k-\sigma+1}{2fk\sigma}) \\
w_j &= \tau^{\frac{(\sigma-1)}{\sigma}} \left(\frac{f^x}{f}\right)^{\frac{k+1-\sigma}{k\sigma}} \left[\frac{(1-\alpha)}{2} + \phi^2 \frac{(1+\alpha)}{2}\right]^{\frac{(\sigma-1)}{k\sigma}}, \quad M_j = 0. \qquad (14)
\end{aligned}
$$

We take the total derivatives of the equilibrium conditions and evaluate them at the symmetric equilibrium as stated in (13) to solve for the break point τ_B. To show that the break point τ_B is decreasing in k it is helpful to rearrange (12) to get

$$
\tau_B^{1-\sigma} = \frac{f_x}{f}\left(\frac{f\Theta}{f_x}\right)^{\frac{\sigma-1}{k}} \qquad (15)
$$

The derivative of $\tau_B^{1-\sigma}$ with respect to k is strictly positive:

$$
\begin{aligned}
\frac{\partial \tau_B^{1-\sigma}}{\partial k} =& \frac{f_x}{f}(1-\sigma)\left(\frac{f\Theta}{f_x}\right)^{\frac{\sigma-1}{k}} \log\left(\frac{f\Theta}{f_x}\right) \\
& \frac{2k\alpha\sigma + \sqrt{1+\sigma^2-2\sigma+4k\sigma+4k\sigma^2\left(\alpha^2+\alpha^2 k-\alpha^2\sigma+\sigma-2\right)}}{k^2\sqrt{1+\sigma^2-2\sigma+4k\sigma+4k\sigma^2\left(\alpha^2+\alpha^2 k-\alpha^2\sigma+\sigma-2\right)}} > 0,
\end{aligned}
$$

because from the derivation of (12) it follows that $\Theta = \phi(\tau_B)$ and accordingly, $0 < \Theta < 1$. Moreover, $0 < f/f_x < 1$ such that $\log\left(\frac{f\Theta}{f_x}\right) < 0$. Since $\sigma > 1$, an increase in k has to be accompanied by a decrease in τ in order to restore the equality in (15) . Q.E.D.

The sustain point marks the critical level of trade costs where high-skilled workers are indifferent between leaving a fully agglomerated region and staying in the core. In the following we suppose $\lambda = 1$ (the $\lambda = 0$ scenario can be derived analogously) and determine the level of trade costs where real wages in i equate real wages in j. Accounting for $M_j = 0$ in the price indices from (9) and computing the cutoff productivities for the equilibrium

[18] We could compute analogously the equilibrium for $\lambda = 0$. Note that the total number of firms is greater for the symmetric equilibrium than for agglomeration as long as $1 + \phi < \frac{2(k-\sigma+1)}{1+\alpha}$ while in the standard core periphery model the total number of firms is independent of the regional distribution of economic activity.

values from (14), we can show that the utility differential is given by:

$$\Delta|_{\lambda=1} = P_i^{-\alpha} \left[1 - w_j \left(\frac{f(1-\alpha)}{f^x(1+\alpha)} \right)^{\frac{-\alpha(k+1-\sigma)}{k(1-\sigma)}} \tau^{-\alpha} \left(\frac{k}{k-\sigma+1} \right)^{\frac{-\alpha}{\sigma-1}} \right], \tag{16}$$

where we can plug in w_j from (14). Hence, setting $\Delta|_{\lambda=1} = 0$ provides an implicit solution for the sustain point τ_S.

D Asymmetric regions

In this Appendix, we allow for differences in firm heterogeneity between countries. Most naturally, we interpret this as differences in technology as discussed in Section 5. Several interesting insights unfold with regard to both short-run and long-run agglomeration equilibria. Figure 4 illustrates three combinations of heterogeneity k for three levels of trade costs.

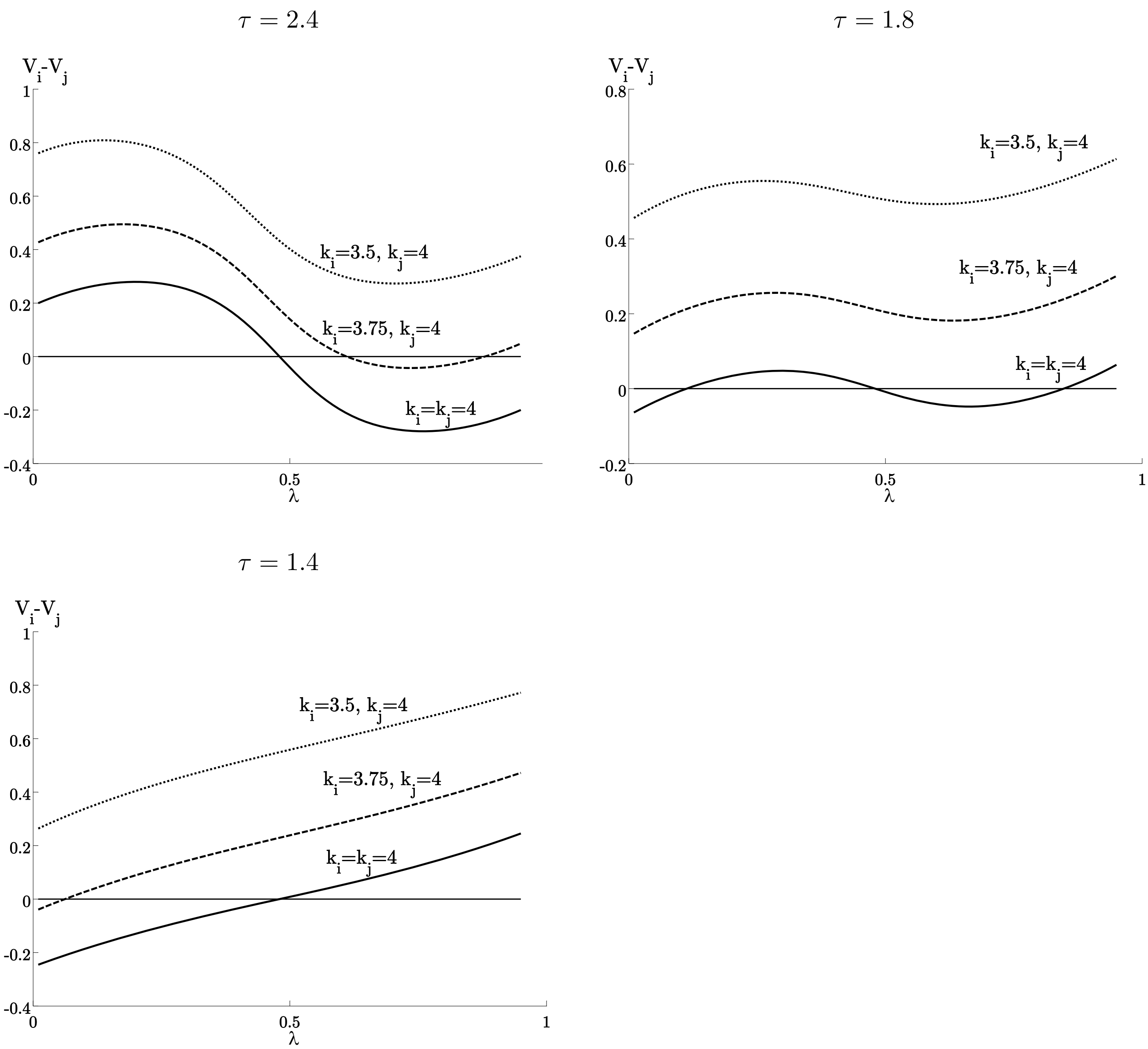
In the high-trade-costs scenario ($\tau = 2.4$) a symmetric productivity distribution $k_i = k_j = 4$ obviously implies a stable equilibrium at $\lambda = 0.5$. Once region i features a productivity distribution that dominates the one in region j according to the hazard rate order, the indirect utility difference $V_i - V_j$ shifts upwards. Hence, for any given allocation of high-skilled laborers the welfare difference between i and j under $k_i = 3.75,\ k_j = 4$ is greater than in the $k_i = k_j = 4$ scenario. The welfare difference is zero only at $\lambda > 0.5$, where the equilibrium at $\lambda \simeq 0.6$ is stable while the interior equilibrium at $\lambda \simeq 0.9$ is unstable (the latter is indicated by the dashed line in Figure 3). Since $V_i - V_j > 0$ at $\lambda = 1 - \varepsilon$ and $V_i - V_j > 0$ for $\lambda = 0 + \varepsilon$ with ε being a small positive number, the agglomeration equilibrium in i is stable while agglomeration in j turns out unstable. Raising the productivity level in i further to $k_i = 3.5$ while keeping k_j fixed implies that interior equilibria do not exist anymore.

In the medium trade costs scenario ($\tau = 1.8$), we observe the welfare difference $V_i - V_j$ being strictly positive for any allocation of high-skilled laborers in both asymmetric scenarios. Accordingly, only full agglomeration in i can occur as a stable equilibrium.

Once the trade cost fall to a sufficiently low level ($\tau = 1.4$), the $\lambda = 0.5$ equilibrium becomes unstable even in the symmetric productivity distribution scenario $k_i = k_j = 4$. Interestingly, if the asymmetry regarding the productivity distributions remains moderate

($k_i = 3.75,\ k_j = 4$) agglomeration in the less productive region j is stable as well as agglomeration in i. This is due to net agglomeration forces being pronounced for very low trade impediments such that the benefit from clustering of economic activity in j may outstrip the productivity advantage in the smaller region i. However, this is a very unlikely equilibrium for two reasons. First, during an integration process full agglomeration of industry in i renders a stable equilibrium for trade costs beyond the level that is necessary for $\lambda = 0$ being a stable equilibrium. Second, a relatively small shock in λ (in Figure 5) $\Delta\lambda \simeq 0.05$ will destabilize the $\lambda = 0$ equilibrium.

Figure 4: Equilibria for asymmetric technologies



References

Baldwin, R.E. and R. Forslid (2000), The core–periphery model and endogenous growth: stabilizing and destabilizing integration, *Economica* **67**, 307–324.

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

Baldwin, R.E. and T. Okubo (2006), Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting, *Journal of Economic Geography* **6**, 323-346.

Behrens, K., G. Mion, and G.I.P. Ottaviano (2011), Economic integration and industry reallocations, in Jovanovic, Miroslav N., (ed.) International Handbook on the Economics of Integration, Volume II, Edward Elgar, Cheltenham, UK.

Caselli, F. (1999), Technological revolutions, *American Economic Review* **89**, 78–102.

Demidova, S. (2008), Productivity improvement and falling trade costs: Boon or bane?, *International Economic Review* **49**, 1437-1462.

Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske (2004), Wage and productivity dispersion in United States manufacturing: The role of computer investment, *Journal of Labor Economics* **22**, 397-429.

Egger, P. and T. Seidel (2008), Agglomeration and fair wages, *Canadian Journal of Economics* **41**, 271-291.

von Ehrlich, M. and T. Seidel (2011), Agglomeration and credit constraints, *unpublished manuscript*, ETH Zurich.

Faggio, G., K. Salvanes, and J. van Reenen (2010), The evolution of inequality in productivity and wages: Panel data evidence, *Industrial and Corporate Change* **19**, 1919–1951.

Falvey, R., D. Greenaway, and Z. Yu (2006), Extending the Melitz model to asymmetric countries, *GEP Research Paper 2006/07.*

Falvey, R., D. Greenaway, and Z. Yu (2011), Catching up or pulling away: Intra-industry trade, productivity gaps and heterogeneous firms, *Open Economies Review* **22**, 17–38.

Helpman, E., M. J. Melitz and S. R. Yeaple (2004), Export versus FDI with heterogeneous firms, *American Economic Review* **94**, 300-316.

Krugman, P. (1980), Scale economies, product differentiation, and the pattern of trade, *American Economic Review* **70**, 950-959.

Krugman, P. (1991), Increasing returns and economic geography, *Journal of Political Economy* **99**, 483-499.

Martin, P. and G.I.P. and Ottaviano (2001), Growth and agglomeration, *International Economic Review* **42**, 947-968.

Melitz, M. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica* **71**, 1695-1725.

Melitz, M. and G.I.P. Ottaviano (2008), Market size, trade, and productivity, *Review of Economic Studies* **75**, 295-316.

Nocke, V. (2006), A gap for me: Entrepreneurs and entry, *Journal of the European Economic Association* **4**, 929-955.

Okubo, T. (2009), Trade liberalisation and agglomeration with firm heterogeneity: Forward and backward linkages, *Regional Science and Urban Economics* **39**, 530-541.

Okubo, T., P.M. Picard and J.-F. Thisse (2010), The spatial selection of heterogeneous firms, *Journal of International Economics* **82**, 230-237.

Ottaviano, G.I.P. (2011), 'New' new economic geography: Firm heterogeneity and agglomeration economies, *Journal of Economic Geography* **11**, 231-240.

Pflüger, M. and J. Südekum (2013), Subsidizing firm entry in open economies, *Journal of Public Economics*, forthcoming.

Saito, H., M. Gopinath and J.J. Wu (2011), Heterogeneous firms, trade liberalization and agglomeration, *Canadian Journal of Economics* **44**, 541-560.

Supplement

A. Constraint on relative wages

In line with the empirical evidence, the heterogeneous-firms literature generally assumes that exporting firms also serve the domestic market. This is equivalent to requiring $\varphi_{ix}^* > \varphi_i^*$ and $\varphi_{jx}^* > \varphi_j^*$ which at the same time ensures that the conditional export probability ranges between zero and unity. In our framework, these assumptions are reflected by the following constraints in the symmetric and asymmetric case:

1. For symmetric countries, $w_i = w_j$ and $\varphi_i^* = \varphi_j^*$, the export cutoffs of the two countries can be stated as functions of the domestic cutoffs:

$$\varphi_{ix}^* = \tau \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}} \varphi_j^* \qquad \varphi_{jx}^* = \tau \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma-1}} \varphi_i^*.$$

Accordingly, for symmetric countries and trade costs converging to unity the export cutoffs are greater than their domestic counterparts as long as $f^x > f$ is satisfied.

2. For asymmetric countries, ensuring that only domestically active firm export restricts the support region of w_i and w_j to

$$\frac{w_i}{w_j} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}} \wedge \frac{w_j}{w_i} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}}$$

$$or$$

$$\frac{w_i}{w_j} < \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}} \wedge \frac{w_j}{w_i} < \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}}, \tag{17}$$

where $\phi \equiv \tau^{-k} \left(f/f_x\right)^{\frac{k-\sigma+1}{\sigma-1}}$. Note that from (5) either the first or the second constraint is binding. In either case the applicable constraint constitutes a necessary and sufficient condition for $\varphi_{ix}^* > \varphi_i^*$ and $\varphi_{jx}^* > \varphi_j^*$ as will be shown in the following.

Think of a situation where $w_i > w_j > 0$, and $\frac{w_i}{w_j} > \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}}$. This implies that the domestic cutoff in j (φ_j^*) as derived in Appendix C is negative while the domestic cutoff in i (φ_i^*) is positive. Recall that $f_x > f$, such that $\varphi_i^* > 0$ and $\varphi_j^* < 0$ imply from equation (4) that the export cutoff in j (φ_{jx}^*) is positive and the export cutoff in i (φ_{ix}^*) is negative. Hence, under the assumptions that $w_i > w_j > 0$ and $\frac{w_i}{w_j} > \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}}$ the ordering of domestic and export cutoffs in region i is $\varphi_{ix}^* < \varphi_i^*$ which means that there are some firms in i that export, but do not produce for the domestic market. Similarly, assuming $w_j > w_i > 0$ and $\frac{w_j}{w_i} > \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}}$ yields $\varphi_{jx}^* < \varphi_j^*$ such that some firms in j produce for the export market only. Accordingly, a necessary assumption for precluding firms in both regions from exporting without producing for the domestic market is $\frac{w_i}{w_j} < \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}} \wedge \frac{w_j}{w_i} < \left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}}$ which ensures that φ_{ix}^* and φ_{jx}^* are positive. However, this condition may not be sufficient as there could be solutions where all cutoff productivities are positive but $\varphi_{ix}^* < \varphi_i^*$ or $\varphi_{jx}^* < \varphi_{jx}^*$ still applies. From equations (4) we observe that $\frac{w_i}{w_j} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}} \wedge \frac{w_j}{w_i} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}}$ has to be fulfilled in such a scenario in order to guarantee $\varphi_{ix}^* > \varphi_i^*$ and $\varphi_{jx}^* > \varphi_j^*$.

Together, the two constraints on relative wages from above constitute sufficient conditions for ensuring that only domestic producers become exporters. Again, consider a situation where $w_i > w_j$, but let the relative wage constraints in (17) be satisfied for now. Then from the domestic-cutoff equation (6), both φ_i^* and φ_j^* are strictly positive. Taking a closer look at equation (6) reveals that – under the above assumptions $\left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}} > \frac{w_i}{w_j}$, and $\left(\frac{1}{\phi}\right)^{\frac{\sigma-1}{\sigma k}} > \frac{w_j}{w_i}$ – the country with the higher wage rate features the lower domestic cutoff productivity. That is $\varphi_j^* > \varphi_i^*$ for $w_i > w_j$ which under the constraints $\frac{w_i}{w_j} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}} \wedge \frac{w_j}{w_i} < \tau^{\frac{\sigma-1}{\sigma}} \left(\frac{f_x}{f}\right)^{\frac{1}{\sigma}}$ implies from equation (4) that $\varphi_{ix}^* > \varphi_{jx}^*$, $\varphi_{ix}^* > \varphi_i^*$, and $\varphi_{jx}^* > \varphi_j^*$ hold true. Therefore, the overall productivity ordering for the $w_i > w_j$ scenario is $\varphi_{ix}^* > \varphi_{jx}^* > \varphi_j^* > \varphi_i^* > 0$. Similarly, the $w_j > w_i$ scenario yields $\varphi_{jx}^* > \varphi_{ix}^* > \varphi_i^* > \varphi_j^* > 0$ under the relative wage constraints in (17).

B. Deriving the break point

The break point is defined by the unique level of trade costs that satisfies (11). At $\lambda = 1/2$ it holds true that $dV_i = -dV_j$. Therefore, it follows that the symmetric equilibrium is stable as long as an additional worker in i decreases real wages in i – which corresponds to an increase of real wages in j and a negative real wage differential $V_i - V_j$. Hence, the break point is characterized by the level of trade costs that satisfies:

$$\frac{dV_i}{d\lambda}|_{\lambda=1/2} = 0 \quad \Leftrightarrow \quad dw_i = \alpha w_i \frac{dP_i}{P_i} \text{ for } \lambda = 1/2 \tag{18}$$

Solving for the break point involves tedious algebra. First, we totally differentiate the labor market clearing condition and trade balance condition. Second, we use this equation system to compute the total derivatives of the price index, the total derivative of the wage rate, and the total derivative of the number of firms (which is part of the price index). Second, we determine the equilibrium values of w_i, P_i, and M_i at $\lambda = 1/2$ and evaluate the above mentioned total derivatives at the symmetric equilibrium.

At the symmetric equilibrium, it can easily be shown from the trade balance condition (8) that the wage is given by $w_i = \frac{\alpha L}{(1-\alpha)H_i}$. To simplify computation we follow most of the literature and normalize the total number of low-skilled workers to $L = 1 - \alpha$ and the total number of high-skilled workers to $H_i + H_j = \alpha$. This implies that the high-skilled wage under symmetry is unity. With $\lambda = 1/2$, $w_i = w_j = 1$, and $M_i = M_j$, the price index from (9) can be simplified to:

$$P_i = M_i^{\frac{1}{1-\sigma}} \left[\left(\frac{\sigma w_i}{(\sigma-1)\,\widetilde{\varphi}_i} \right)^{1-\sigma} + \left(\varphi_j^*/\varphi_{jx}^*\right)^k \left(\frac{\tau\sigma w_j}{(\sigma-1)\,\widetilde{\varphi}_{jx}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

$$= M_i^{\frac{1}{1-\sigma}} \frac{\sigma}{(\sigma-1)} \left[\left(\frac{1}{\widetilde{\varphi}_i} \right)^{1-\sigma} + \left(\varphi_j^*/\varphi_{jx}^*\right)^k \left(\frac{\tau}{\widetilde{\varphi}_{jx}} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{19}$$

Using the price index jointly with the labor-market constraint (7) we can compute the equilibrium number of firms under symmetry:

$$M_i = \frac{\alpha\,(1+k-\sigma)}{2f\,(1+\phi)\,k\sigma}. \tag{20}$$

Using the total derivatives of the labor market constraint (7) together with the total differential of the trade balance condition (8) – both evaluated at $\lambda = 1/2$ – we can compute:

$$\frac{dw_i}{d\lambda} = -\frac{2\,(\phi-1)\,(\sigma-1)\,[\phi(1+\alpha)-(1-\alpha)]}{(1-\alpha)(\sigma-1)+2\phi\,(1+2k\sigma-\sigma)+(1+\alpha)\,(\sigma-1)\,\phi^2}, \tag{21}$$

and

$$\frac{dM_i}{d\lambda} = \frac{\alpha\,(1+k-\sigma)\left[(1-\alpha)\,(\sigma-1)+2\phi+2\phi\sigma\,(k-1+k\alpha)+(1+\alpha)\,(\sigma-1+2k\sigma)\,\phi^2\right]}{fk\sigma\left[(1-\alpha)\,(\sigma-1)+2\phi+2\phi\sigma\,(2k-1)+(1+\alpha)\,(\sigma-1)\,\phi^2\right]} \tag{22}$$

Note that $\frac{dw_i}{d\lambda}$ may be positive or negative depending on whether $\phi \lesseqgtr \frac{(1-\alpha)}{(1+\alpha)}$ while it is easy to show that $\frac{dM_i}{d\lambda} > 0$ always holds true. As in the classical core-periphery model, the level of trade freeness determines whether wages increase or decrease as a response to migration of high-skilled laborers from j to i. In other words, either the competition effect or the home-market effect dominates.

Finally, we need to compute the total derivative of the price index and evaluate it at $\lambda = 1/2$. This yields:

$$\frac{dP_i}{P_i} = \frac{(\phi-1)^2}{M_i\,(\phi^2-1)\,(\sigma-1)}dM_i - \frac{(\phi-1)^2 - 2\sigma\left(1+\phi k+\phi^2+\phi^2 k\right)+\sigma^2\,(1+\phi)^2}{(\phi^2-1)\,(\sigma-1)^2}dw_i \tag{23}$$

$$= \frac{2fk\,(\phi-1)\,\sigma}{\alpha\,(1+k-\sigma)\,(\sigma-1)}dM_i - \frac{(\phi-1)^2 - 2\sigma\left(1+\phi k+\phi^2+\phi^2 k\right)+\sigma^2\,(1+\phi)^2}{(\phi^2-1)\,(\sigma-1)^2}dw_i$$

Plugging (22), (21), and (23) into the definition of the break point from (18) we compute the critical level of trade freeness τ_B for which the symmetric equilibrium becomes unstable as stated in (12).