Localized technological externalities and the geographical distribution of firms

Giulio Bottazzi[1,2] and Pietro Dindo[*,1]

[1] *CAFED and LEM, Scuola Superiore Sant'Anna*
Piazza Martiri della Libertà 33, 56127 Pisa, Italy
[2] *BRICK, Collegio Carlo Alberto*
Via Real Collegio, 30, 10024 Moncalieri (Torino), Italy

February 1, 2010

Abstract

Using an analytically solvable equilibrium model, we study how the distribution of economic activities is affected by the trade-off between pecuniary externalities, as dependent on transportation costs, and localized technological externalities, as dependent on inter-regional spillovers. We model localized technological externalities as having a cost saving effect that can be interpreted both as a tangible technological advantage, like the presence of a stronger industrial infrastructure, and as an intangible advantage, like a more efficient labour force composition or the presence of some inter-firms knowledge spillover. Under the assumption of capital mobility and labour immobility, we show that whereas decreasing transportation costs, i.e. promoting market openness, leads to sudden agglomeration, increasing inter-regional spillovers, i.e. promoting technological openness, favors a smoother transition between different levels of firms concentration and ultimately leads to a less uneven distribution of welfare.

Keywords: New Economic Geography; Agglomeration; Footloose capital models; Technological externalities; Market and technological openness

JEL Classification: F12, F15, R12, O3

1 Introduction

The skewed nature of the distribution of economic activities found in both developed and developing countries, at any scale, from cities to worldwide geographical areas, can be the result of both market mediated interactions, such as labor pooling or intermediate goods availability, and non tradable differences across the geographical locations (c.f. the three sources of agglomeration in Marshall, 1920). Beside the effect of trade openness in final markets and increased mobility in factors of productions, like labour and capital, economic agglomeration is plausibly enhanced by the institutional framework, the availability of public

*Tel.: +39-050-883341; fax: +39-050-883344; e-mail: p.dindo@sssup.it.

infrastructures, higher levels of human and social capital and the local and tacit nature of technical knowledge. Indeed, the abundant presence of agglomerated production clusters away from big cities and main transport systems suggests that forces other than transportation costs, advantages due to larger local demand, or deeper factor markets, are at work. These forces are not exclusively acting in high-tech sectors, like semiconductors or ICT services, but are often pervasive of the entire economy. Analyzing Italian manufacturing industry, Bottazzi et al. (2008) find that sectors like Food Products, Leather Products or Basic Metal Workings are highly agglomerated and their agglomeration cannot be explained by the presence of transport infrastructures or localized demand. This is not a peculiar aspect of Italian manufacturing, as similar results have been found for the US (Ellison and Glaeser, 1997), France (Maurel and Sedillot, 1999), Germany (Brenner, 2006) and the UK (Devereux et al., 2004).

If localized non-pecuniary advantages are important in describing the observed final outcome of firms locational choice, at least as much as pecuniary market-mediated interactions, it becomes relevant to investigate what aggregate effects can be observed when the political, cultural and social barriers which make these advantages local are, at least partially, abated. Indeed there are several local factors which affect productions and its efficiency that could be, in principle, made more global by the enactment of suitable policies. Think for instance of the appropriate management of technological innovation, which can be improved by common commercial laws, of labour-embedded technological capabilities, which can be shared through common education, or of the adoption of global industrial standards. How does a variation in the degree of "openness" of these social and technological factors and, in particular, their interaction with the freeness of trade and the commercial "globalization", affect the geographical distribution of firms?

In the present paper we intend to address this question inside the domain of New Economic Geography (NEG). Since Krugman (1991b) this literature has mostly dealt with the effect of pecuniary externalities on the spatial distribution of economic activities. Firms agglomeration arises as the result of a market mediated circular causation: firms locate where demand is high and demand moves where there are many firms. In the original models, localized technological externalities were disposed off explicitly as sources of economic agglomeration, essentially because presumed to be too difficult to properly measure and too easy to model (Krugman, 1991a, p.53). Notwithstanding the original lack of interest, recent theoretical contributions to NEG have extended the investigation by including non-pecuniary external economies. A number of works have adopted the type of externality introduced in growth models by Grossman and Helpman (1980), postulating an R&D sector with marginal cost of innovation decreasing in the number of existing innovations. In particular Martin (1999) and Martin and Ottaviano (1999) include a Grossman-Helpman type of externality in a footloose capital model where capital is mobile and labour is not. Despite being the endogenous force behind growth, the technological externality is not causing agglomeration because its advantage can be globally exploited. In fact, patents or capital created a lower cost in the most innovative region can migrate and exert their innovativeness also in the other region. This is not the case of Baldwin and Forslid (2000), who consider the same type of technological externality, but this time with mobile workers and immobile capital. In their model each location has its own R&D sector so that the technological externality can be a source of agglomeration. Moreover they introduce an inter-regional spillover parameter, a sort of "technological openness", which measures to what extent non-pecuniary advantages are location specific or can be shared across regions. Increasing the flow of knowledge between locations reduces innovation costs in all regions, so that location choice is less relevant and agglomerated outcomes less likely. Ultimately, the equilibria of the economy and their desirability in terms of welfare are decided by the interplay between market openness, as dependent on trade costs, and technological openness, as dependent on inter-regional spillovers. The model is however not analytically tractable, and the authors analyze stability only for a predetermined set of benchmark equilibria corresponding to full agglomerated and symmetrically non-agglomerated economies.

In the present paper we advance an analytically tractable general equilibrium model that explicitly accounts for the presence of technological externalities. Following Forslid and Ottaviano (2003), we obtain analytical tractability through partial factor immobility. In particular, we impose labor immobility and assume that households are both local workers and global investors, as in Martin and Rogers (1995). In this way the mobile factor is represented by the capital, whose rent is payed to households/shareholders and consumed in the location in which they reside. We believe that this assumption better represents today increased capital mobility, specially in geographical area like the European Union, in which relative regional homogeneity leads to flows of capital which are hardly matched by flow in any other productive factor. While Martin and Rogers (1995) and the more recent Dupont and Martin (2006), lacking any self-reinforcing mechanism, reproduce firms agglomeration exclusively via the so called home market effect, when there are local differences in the exogenous endowment of factors, in our model agglomeration can emerge also with a priori identical locations and is sustained by the endogenous effect of the technological externality. This externality is introduced through a mechanism inspired by Grossman and Helpman (1980). The enhancement of capital creation capability due to scale economies in R&D activity is modeled as a direct effect on final good producers by an increase in operating margins due to a sharing of fixed costs. Firms might share costs of the maintenance of public services and infrastructures, or of labour training, thus making the distinction between modern sector and traditional sector workers endogenous. Our cost sharing assumption can be interpreted both as a tangible technological advantage, like the presence of a stronger industrial infrastructure, and as an intangible advantage, like a more efficient labour force composition or the presence of some inter-firms knowledge spillover. Inside this framework, we introduce an inter-regional knowledge spillover parameter describing the degree of localization of the technological externality, as in Baldwin and Forslid (2000). The analytical tractability of our model allows for the explicit derivation of geographical equilibria, defined as those distribution of firms where households do not have incentives to change capital allocation.

Our analysis confirms previous findings by Baldwin and Forslid (2000) about the stabilizing nature of knowledge spillovers: the higher the spillover the larger the interval of transportation costs which lead to firms equidistribution. We also find that if the spillover is high enough, there exists a smooth equilibrium transition between agglomeration and equidistribution, with partly agglomerated economy for intermediate values of transportation costs. In this case an opening of inter-regional trade does not entail an abrupt reallocation of economic activities nor the hysteresis effect, typical of NEG model, which locks the economy in a core-periphery equilibrium also if higher trade costs are reintroduced. Welfare analysis reveals that agglomeration can entail lower welfare level for the economic periphery, also when it represents the geographic equilibrium. The huge "welfare gap" existing between core and periphery regions, which often hinders the implementation of trade opening policies, could be reduced and eventually eliminated if market and technological integrations were pursued together. We provide conditions under which either policies are to be preferred.

The rest of the paper is organized as follows. In Section 2 we introduce the model and derive the market equilibrium. In Section 3 we find the geographical equilibria of our economy by analyzing how changes in the distribution of capital influence each region capital rents. In Section 4 we complete the characterization of geographical equilibria by making explicit their dependence on the parameters ruling trade openness and technological openness. The welfare analysis is presented in Section 5. Section 6 concludes.

2 The model

Consider an economy with two locations. Each location is populated by L households who are endowed with labour and capital and supply them inelastically. The economy has a modern and a traditional sector. Whereas the traditional sector supplies an homogeneous good, the modern sector supplies differentiated products. In both sectors production is localized.

Households are "local" workers and "global" consumers, that is, they can buy goods produced in both locations and traded in a global market. At the same time, households are "global" investors, that is, they can supply capital to both locations. Modern goods are traded at a transportation cost which takes the form of an iceberg cost: for each unit of good shipped, only a fraction $\tau \in (0,1]$ arrives at destination. τ is thus an inverse index of transportation costs, or an index of freeness of trade. Traditional goods and capital are traded at no costs.

Consumption All households have the same preferences and decide how much of the traditional good C_T and of the bundle of modern goods C_M to consume so as to maximize a Cobb-Douglas utility function

$$\mathrm{U} = \mathrm{C}_T^{1-\mu}\mathrm{C}_M^{\mu}\,, \tag{2.1}$$

with $\mu \in (0,1)$. As a result a fraction μ of each household income is spent on C_M and a fraction $(1-\mu)$ is spent on C_T. The utility of the bundle C_M is of constant elasticity of substitution (CES) type,

$$\mathrm{C}_M = \left(\sum_{i=1,N} c_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)}} \qquad \sigma > 1\,, \tag{2.2}$$

with c_i the consumption of good i, $i = 1,\ldots,N$. This implies that the N modern goods are substitutes, with a mutual elasticity of substitution equal to σ (cfr. Dixit and Stiglitz, 1977).

Production Each household is endowed with one unit of labour, and there is not an *a priori* distinction between workers of the modern and traditional sector. The traditional sector uses only labour as input under constant returns to scale with unitary marginal costs. Due to the large number of potential producers, $2L(1-\mu)$ at equilibrium, this market is perfectly competitive and the traditional good is sold at its marginal cost, which we take as the normalization price of the economy.

We assume a fixed number of firms, N, active in the modern sector. Both capital and labour are used in the production of modern goods. The amount of labour v_i that firm i employs to produce an amount y_i of modern output is given by the usual scale economy cost function

$$v_i = \beta y_i + \alpha_{l_i}\,, \tag{2.3}$$

where β is constant across firms and across locations and, in our case, the fixed amount of labour to start production, α_{l_i}, might depend on the location l_i of firm i. Each firm also needs one unit of capital, available at a price r_i. This, at equilibrium, is given by the operating profits

$$r_i = p_i y_i - w_i v_i \,, \tag{2.4}$$

where p_i is the price of good i and w_i is the cost of labour of firm i.

Given the structure of preferences in (2.2) each firm produces a different product. The total number of varieties produced in each location is thus equal to the amount of capital available there.

We assume that each household is endowed with the same amount of capital $N/2L$. Since they are interested in maximizing the capital revenues, and since capital is moved without costs, their investment choices are symmetric so that each household invest a fraction $1/2L$ of capital in each firm..[1]

The market structure is that of monopolistic competition, that is, each firm maximizes its profits given market demand elasticity and irrespectively of other firms behavior.

Technological externality So far our assumptions closely mimic footloose capital models, such as Martin and Rogers (1995) or Dupont and Martin (2006). Departing from these works, we introduce a localized technological externality, which we model as a term of direct firms interaction, not mediated by market forces. More specifically we assume that firms labour fixed cost α_l decreases with the number of firms located in each region according to

$$\alpha_l = \frac{N\alpha}{n_l + \lambda(N - n_l)} \,, \tag{2.5}$$

where n_l is the number of firms producing modern goods in location $l \in \{1, 2\}$. Equation (2.5) represents a positive localized externality because the local labour fixed costs are decreasing in the number of firms. The parameter $\lambda \in (0, 1]$ governs the inter-regional spillover, that is, the degree to which technological externalities are localized, thus accounting for the cost-reducing effect of firms in the other region.

Equation (2.5) is analogous to the the production function of the R&D sector as in Grossman and Helpman (1980), used also in a geographic context by Baldwin and Forslid (2000). The latter work even consider the impact of a inter-regional spillover λ as we do here. Our interpretation is, however, slightly different, as we do not model a R&D sector and a market for patents, but instead assume that the externality operates as a fixed cost reduction.

In fact, equation (2.5) can be tought as firms sharing training costs of creating human capital, for instance by transforming an endogenously determined share of unskilled workers of the traditional sector in skilled workers of the modern sector. In this case λ represents the inter-regional knowledge spillover and measures the degree of integration in training or the cultural and technical linkages between the two regions.

Alternatively, (2.5) can be interpreted as a form of cost sharing for setting up public infrastructures which, by lowering the total fixed costs, facilitates the production activities. Firms might share the costs of maintaining a distribution network, commercial or legal services, or whatever they might collectively do to enhance production. In particular (2.5) can capture the advantages deriving from a research sector whose efficiency increases with the number of

[1]This is what it is usually assumed in footloose capital models (see Baldwin et al. (2003) p.74) to avoid the complications resulting from household strategic interaction. This assumption is harmless at a geographical equilibrium, that is, at a distribution of firms where either rents in both locations are equal or all firms are in the same location. It is, however, not harmless out of equilibrium and stability results might depend on it.

active firms, as modeled explicitly in Martin (1999); Martin and Ottaviano (1999); Baldwin and Forslid (2000)

Ultimately the cost saving effect can be interpreted both as a tangible technological advantage, like the presence of a stronger industrial infrastructure, and as an intangible advantage, like a more efficient labour force composition or the presence of some inter-firms knowledge spillover.

The inter-regional spillover parameter captures the "technological openness" of the economy: for low values of λ the economy is split in technological separated parts and firms can exclusively exploit advantages derived by co-location in the same region. For high values of λ, conversely, the economy is technologically integrated, and production-improving externalities are also allowed to flow across regions. Increasing the flow of knowledge between locations has two effects. First, it creates a global advantage in reducing production cost of all firms, and second, it makes location choice less relevant and uneven outcomes less likely.

2.1 Market equilibria

Having specified all the elements of our economy, we derive, for any given fixed distribution of firms, equilibrium capital rents for each of the two locations. Due to perfect competition and constant returns to scale in the traditional sector, here wages are equal to prices. Moreover, due to zero transportation costs prices, and thus wages, must be the same in both locations.

When consumers are not mobile, if the economy is at an equilibrium, it should be indifferent for a worker to work in the traditional or modern sector. As a result wages in the two sectors, and in the two locations, are equal. For this reason it is convenient to use wages to normalize prices in the economy.

In order to find equilibrium prices, quantities, and profits of the modern sector, one should in principle analyze each of the N product markets. Nevertheless the problem can be simplified by considering only a representative market for each location. In fact, location by location, firms produce using the same technology, face the same demand (due to the CES utility all goods are substitutes), and the same labour supply. This implies that equilibrium prices, quantities and wages are the same for all the firms in a given location. We can thus consider only two representative product markets, one for each location l.

We proceed as follows. Exploiting the CES preference structure (2.2), we compute consumer demand for the goods produced in each location, taking into account that all goods are substitutes, that transportation costs impact the consumption of foreign goods, and that the budget constraint depends also on the rent of capital. Using the monopolistic competition structure of the market, and knowing consumers demand elasticity, we derive firms pricing behavior. By setting supply equal to demand we are able to determine equilibrium quantities and capital rents as a function of the parameters of the economy and the distribution of firms across locations. The following step is to use capital rents and labour income to determine consumers demand for the traditional goods. Since this market is at equilibrium, we can derive the total labour demand. The last step will be to check that the labour market is at equilibrium too. Since there are two segmented labour markets, requiring that both clear amounts to posing a constraint on agents preferences and on the scale of the economy.

Let us start from consumers demands. Denote the quantity consumed by a consumer who resides in l of a product produced in m as d_{lm} with $l.m \in \{1, 2\}$. Relative demand under CES utility satisfies

$$\frac{d_{11}}{d_{12}} = \left(\frac{p_2}{p_1\tau}\right)^{\sigma} \quad \text{and} \quad \frac{d_{22}}{d_{21}} = \left(\frac{p_1}{p_2\tau}\right)^{\sigma} , \tag{2.6}$$

while agents budget constraint gives

$$\begin{cases} \mu I(n_1, n_2) = n_1 d_{11} p_1 + n_2 d_{12} \dfrac{p_2}{\tau} \\ \mu I(n_1, n_2) = n_1 d_{21} \dfrac{p_1}{\tau} + n_2 d_{22} p_2 \, , \end{cases} \tag{2.7}$$

and $I(n_1, n_2)$ is the income of each consumer given by his wage, set to 1, plus his share of capital rent as depending on the geographical distribution of firms (n_1, n_2). Using the previous equations to find demands leads to

$$\begin{aligned} d_{11} &= \frac{\mu I(n_1, n_2)}{n_1 p_1 + n_2 p_1^{\sigma} p_2^{1-\sigma} \tau^{\sigma-1}} \qquad & d_{12} &= \frac{\mu I(n_1, n_2)\tau^{\sigma}}{n_1 p_1^{1-\sigma} p_2^{\sigma} + n_2 p_2 \tau^{\sigma-1}} \\ d_{22} &= \frac{\mu I(n_1, n_2)}{n_1 p_1^{1-\sigma} p_2^{\sigma} \tau^{\sigma-1} + n_2 p_2} \qquad & d_{21} &= \frac{\mu I(n_1, n_2)\tau^{\sigma}}{n_1 p_1 \tau^{\sigma-1} + n_2 p_1^{\sigma} p_2^{1-\sigma}} \, . \end{aligned} \tag{2.8}$$

Given the market structure of monopolistic competition, each firm, knowing consumers inverse demand, sets the output so that marginal revenues are equal to marginal costs. In location l this gives

$$p_l \left(1 + \frac{1}{\varepsilon}\right) = \beta \, , \tag{2.9}$$

where $\varepsilon = \partial \log c / \partial \log p$ is the demand elasticity, and we have used the fact that wages are normalized to one. Given (2.2), as long as the number of commodities N is large (see Dixit and Stiglitz, 1977, for the details), it holds that

$$\varepsilon = -\sigma,$$

which together with (2.9) implies

$$p_l = \beta \frac{\sigma}{\sigma - 1} \, . \tag{2.10}$$

Equating, location by location, demand and supply, we get

$$\begin{cases} y_1 &= \quad L d_{11} + L \dfrac{d_{21}}{\tau} \\ y_2 &= \quad L \dfrac{d_{12}}{\tau} + L d_{22} \, , \end{cases} \tag{2.11}$$

where, due to iceberg costs, only a fraction τ of the imported output is consumed. Using the demand derived in (2.8) and substituting the expression for prices in (2.10) we can easily solve for market equilibrium quantities

$$\begin{cases} y_1 &= \quad \mu I(n_1, n_2) L \, \dfrac{\sigma - 1}{\beta\sigma} \left(\dfrac{1}{n_1 + n_2 \tau^{\sigma-1}} + \dfrac{\tau^{\sigma-1}}{n_1 \tau^{\sigma-1} + n_2} \right) \\ \\ y_2 &= \quad \mu I(n_1, n_2) L \, \dfrac{\sigma - 1}{\beta\sigma} \left(\dfrac{\tau^{\sigma-1}}{n_2 \tau^{\sigma-1} + n_1} + \dfrac{1}{n_2 + n_1 \tau^{\sigma-1}} \right) \, . \end{cases} \tag{2.12}$$

The revenue of a firm in l is now given by

$$p_l y_l = \beta \frac{\sigma}{\sigma - 1} y_l \,, \tag{2.13}$$

and, using (2.4), its capital rent is

$$r_l = \frac{\beta}{\sigma - 1} y_l - \alpha_l \,. \tag{2.14}$$

Introducing the fraction of firms (or capital) in location 1, $x = n_1/N$ (so that $n_2 = (1-x)N$), rents payed by each firm can be finally written as a function of x

$$\begin{cases} r_1(x) &= \dfrac{\mu I(x) L}{N\sigma} \left(\dfrac{1}{x + (1-x)\tau^{\sigma-1}} + \dfrac{\tau^{\sigma-1}}{x\tau^{\sigma-1} + (1-x)} \right) - \dfrac{\alpha}{x + \lambda(1-x)} \\ r_2(x) &= \dfrac{\mu I(x) L}{N\sigma} \left(\dfrac{1}{x\tau^{\sigma-1} + (1-x)} + \dfrac{\tau^{\sigma-1}}{x + (1-x)\tau^{\sigma-1}} \right) - \dfrac{\alpha}{1 - x + \lambda x} \,, \end{cases} \tag{2.15}$$

where we have used (2.5) for α_l. Notice that this is still an implicit equation because $I(x)$ is a function of r_1 and r_2. In fact, with wages normalized to one, and remembering that each households invests a fraction $1/2L$ of capital in each firm, we have

$$I(x) = 1 + \frac{N}{2L} (x r_1(x) + (1-x) r_2(x)) \,. \tag{2.16}$$

Solving (2.15) for $r_1(x)$ and $r_2(x)$ we get

$$\begin{cases} r_1(x) = \left(1 + \dfrac{R(x)}{2L}\right) \dfrac{\mu L}{N\sigma} \left(\dfrac{1}{x + (1-x)\tau^{\sigma-1}} + \dfrac{\tau^{\sigma-1}}{x\tau^{\sigma-1} + (1-x)} \right) - \dfrac{\alpha}{x + \lambda(1-x)} \\ r_2(x) = \left(1 + \dfrac{R(x)}{2L}\right) \dfrac{\mu L}{N\sigma} \left(\dfrac{1}{x\tau^{\sigma-1} + (1-x)} + \dfrac{\tau^{\sigma-1}}{x + (1-x)\tau^{\sigma-1}} \right) - \dfrac{\alpha}{1 - x + \lambda x} \,, \end{cases}$$

where $R(x) = N(x r_1(x) + (1-x) r_2(x))$ is the total rent of the economy. It remains to be checked that, location by location, firms' labour demand is always smaller than L. This will be done in the following where we summarize our findings

Proposition 2.1. *For any geographical distribution* (n_1, n_2)*, provided that*

$$S = \frac{N}{L} < \frac{\mu + \sigma - 2\mu\sigma}{\alpha\sigma(1-\mu)} = \tilde{S} \,, \tag{2.17}$$

which requires that $\tilde{S} > 0$*, there always exists a unique level of prices and wages such that the global traditional market, the* N *global modern markets, and the two local labour markets are at equilibrium. All firms in the same location pay the same capital rent which, as a function of* $x = n_1/N$*, is given by*[2]

$$\begin{cases} r_1(x) = \left(1 + \dfrac{R(x)}{2L}\right) \dfrac{\mu L}{N\sigma} \left(\dfrac{1}{x + (1-x)\tau^{\sigma-1}} + \dfrac{\tau^{\sigma-1}}{x\tau^{\sigma-1} + (1-x)} \right) - \dfrac{\alpha}{x + \lambda(1-x)} \\ r_2(x) = \left(1 + \dfrac{R(x)}{2L}\right) \dfrac{\mu L}{N\sigma} \left(\dfrac{1}{x\tau^{\sigma-1} + (1-x)} + \dfrac{\tau^{\sigma-1}}{x + (1-x)\tau^{\sigma-1}} \right) - \dfrac{\alpha}{(1-x) + \lambda x} \,, \end{cases} \tag{2.18}$$

[2]Throughout the paper, without loss of generality, we consider x to be a real number in the interval $[0, 1]$.

where the total capital rent of the economy reads

$$R(x) = N(xr_1(x) + (1-x)r_2(x)) = \frac{2\mu L - N\alpha\sigma\left(\frac{x}{x+\lambda(1-x)} + \frac{1-x}{(1-x)+\lambda x}\right)}{\sigma - \mu}. \tag{2.19}$$

Proof. The proof that capital rents are as in (2.18) is in the text. Condition (2.17) can be found by imposing that the maximal demand for labour employed in the modern sector is lower than L. In particular, $\tilde{S}$ is the value of N/L such that the demand for labour employed in the modern sector, when fully agglomerated in one of the two regions, is exactly L. □

Due to the segmentation of the labour market, the economy is at an equilibrium only provided that the ratio of the number of firms to the number of households is smaller than a given threshold $\tilde{S}$. Obviously the foregoing condition can only be met when $\tilde{S} > 0$, or, in terms of the preference for the modern goods, $\mu < \sigma/(2\sigma - 1)$.[3] Summing up, for any given elasticity of substitution, provided that preferences for modern goods are not too strong, there always exists a range of firms-to-households ratios such that markets in both location are at equilibrium and capital rents are given by (2.18).

The dependence of equilibrium capital rents on the geographical distribution of firms is due to both pecuniary and technological externalities. The effect of the former goes via the sum of local and foreign demand, that is the part of capital rents in (2.18) which depends on τ. When concentration of local firms is low, each firm faces a high local demand and makes high profits. As the concentration of local firms increases, a higher competition lowers the profits coming from the local demand which, due to positive transportation costs, are not fully compensated by an increased foreign demand. Thus pecuniary externalities have a negative effect on agglomeration. The effect becomes stronger as transportation costs increase, that is, for lower τ.

Technological externalities influence equilibrium capital rents both locally and globally. The local effect is due to the direct dependence of firm fixed costs on the geographical distribution of firms, as given by the last term of both expressions in (2.18). The higher the concentration of firms in location l, the lower the fixed costs payed by the industry there, and the higher the capital rent in l. The global effect of the technological externality is due to the dependence of local rents in (2.18) on global capital rents $R(x)$ and acts as a sort of multiplier. In fact, the geographical distribution has first an impact on total fixed costs, and thus on total capital rents $R(x)$, which, in turn, have a wealth effect on consumers demand, thus affecting each location capital rents. An increase in the concentration of firms, lowers total fixed costs payed by all firms, increases total capital rents, increases households wealth, increases total demand and, in turn, increases capital rents further in a multiplier fashion. Notice that whereas the local effect increases a given location rents only under an increase of firms agglomeration in that given location, the global effect increases capital rents in both locations, no matter where firms do actually agglomerate. As we shall see the overall effect of these two forces and its strength depend both on the inter-regional spillover λ and on the freeness of trade τ. In general when λ is low (high) the technological externality is (not) localized and its variability with the local concentration is high (low). In the extreme case, $\lambda = 1$, both regions have equal benefits, irrespectively of the geographical distribution of firms.

[3]This is the same condition on consumers preferences found in e.g. Forslid and Ottaviano (2003). Whereas in that paper the long run no-entry condition (zero profit) fixes the ratio S to its maximum size $\tilde{S}$, here we assume that the ratio S is exogenously given and allow capital rents to be positive.

3 Geographical equilibria

A distributions x of capital, and thus firms, is a geographical equilibrium of the economy when, in the search for higher rents, firms have no incentives to move production (or, equivalently, households have no incentives to move capital) from one location to the other. Geographical equilibria can be of two types. At a *border* equilibrium capital is concentrated in one location, say 1 so that $x = 1$, and rents in 1 are higher than rents in 2. At an *interior* equilibrium, capital is distributed between the two locations, that is $x \in (0,1)$, rents are equal, and moving capital from one location to the other lowers its rent. In what follows we shall name the border equilibria as agglomerated economies (AG) and the interior equilibria either as partially agglomerated economies (PAG), when $x \neq 1/2$, or as non-agglomerated economies (NAG), when $x = 1/2$.

For what concerns the present analysis we shall assume that, for any value of the inter-regional spillover λ and of the freeness of trade τ, the total rent of the economy is positive. It turns out that the positive rent condition can be written as another constraint on the firms to households ratio S.

Lemma 3.1. *Consider capital rents* $r_1(x)$ *and* $r_2(x)$ *as in* (2.18)*. The total rent of the economy,* $R(x) = N(xr_1(x) + (1-x)r_2(x))$ *is positive for any value of the transportation cost* τ *and of the inter-regional spillover* λ *provided that*

$$S = \frac{N}{L} < \frac{\mu}{\alpha\sigma} = \bar{S}\,. \tag{3.1}$$

Proof. By computing the total rent profits one immediately sees that they do not depend on τ but do depend on λ. The given value of $\bar{S}$ has been found by equating to zero the minimum values of total rents (2.19), obtained when $\lambda = 0$. □

In what follows we shall assume that the scale S of the economy is such that both constraints in (2.17) and (3.1) hold. Since it is always possible to order the two values of $\tilde{S}$ and $\bar{S}$, this is equivalent to assume that $S < \min\{\tilde{S}, \bar{S}\}$.[4]

We can establish the existence of geographical equilibria by looking at the rent difference $\Delta(x) = r_1(x) - r_2(x)$. A geographical distribution $x \in (0,1)$ is an interior equilibrium provided that $\Delta(x) = 0$ and $\frac{d\Delta(x)}{dx} = \Delta'(x) < 0$. In fact, since incentives to move capital and destabilize an interior distribution x are present only when the rent difference is increasing in x, negative (positive) $\Delta'(x)$ means that x is (not) an interior equilibrium. The two border equilibria $x = 0$ and $x = 1$ are instead characterized, respectively, by $\Delta(0) < 0$ and $\Delta(1) > 0$. A negative (positive) $\Delta(0)$ means that when the production of modern goods is concentrated in location 2, capital rents are higher (lower) there than in 1 so that $x = 0$ is (not) an equilibrium. The same argument, with the reversed sign, holds for $x = 1$.

Let us start from the existence of geographical equilibria corresponding to NAG and AG. Given the symmetry of the economy it always holds $\Delta(1/2) = 0$ so that NAG is an equilibrium whenever $\Delta'(1/2)$ is negative. The same symmetry also implies that the sign of $\Delta(0)$ is the opposite of the sign of $\Delta(1)$, so that only one of the two signs needs to be checked to prove that AG is an equilibrium. As a result the combinations of the signs of $\Delta(0)$ and $\Delta'(1/2)$ completely characterize whether the AG and/or NAG configurations appear. When $\Delta(0)$ is

[4]By simple computations one can show that $\bar{S} = \min\{\tilde{S}, \bar{S}\}$ when $\mu \in [0,1/2]$ or $\mu \in (1/2,1]$ and $\sigma < \mu^2/(2\mu - 1)$ and $\tilde{S} = \min\{\tilde{S}, \bar{S}\}$ otherwise.

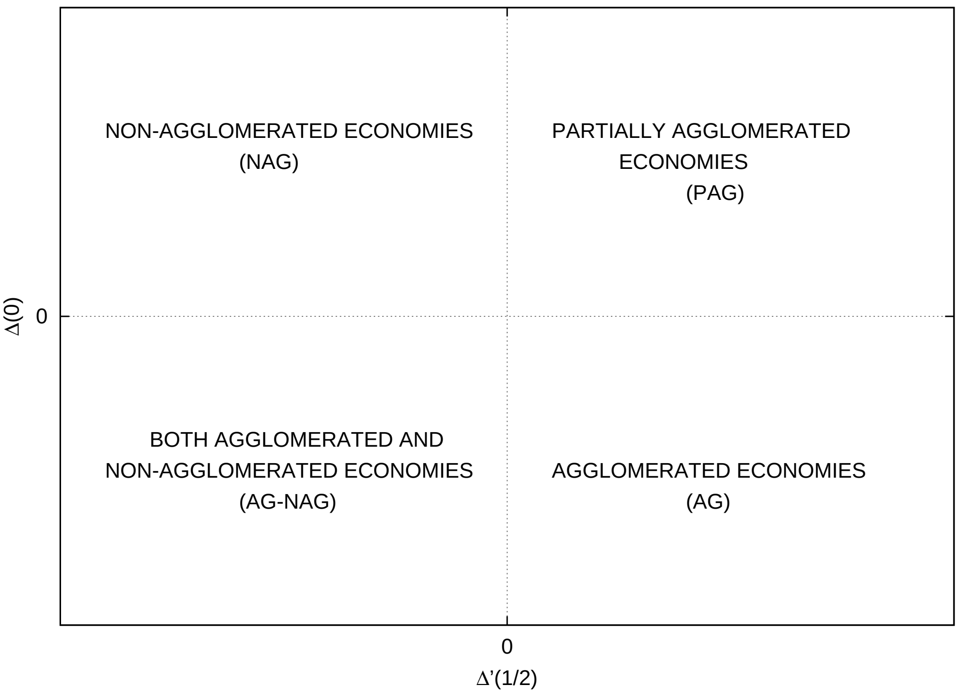

Figure 1: Existence of different type of geographical equilibria in the plane $(\Delta'(1/2), \Delta(0))$.

negative and $\Delta'(1/2)$ is positive, the shape of the rent differential $\Delta(x)$ is such that AG is a geographical equilibrium whereas NAG is not. When $\Delta(0)$ and $\Delta'(1/2)$ are both negative, AG and NAG are both equilibria. When $\Delta(0)$ is positive and $\Delta'(1/2)$ negative, NAG is an equilibrium whereas AG is not. When $\Delta(0)$ and $\Delta'(1/2)$ are both positive, neither AG nor NAG are equilibria.

It remains to be seen whether there exist interior equilibria corresponding to PAG. It turns out that the knowledge of the signs of $\Delta(0)$ and $\Delta'(1/2)$ is enough to answer this question too. This will be shown in the proof of the following proposition where a formal characterization of all the geographical equilibria is given. These results are summarized in Fig. 1. Figures 2 gives instead specific examples of the functional forms of the rent difference $\Delta(x)$.

Proposition 3.1. *Consider the capital rent difference $\Delta(x)$ and assume $\tilde{S} > 0$ and $S < min\{\tilde{S}, \bar{S}\}$. It holds that*

- *if $\Delta'(1/2) > 0$, $\Delta(0) < 0$, only AG is a geographical equilibrium,*
- *if $\Delta'(1/2) < 0$, $\Delta(0) > 0$, only NAG is a geographical equilibrium,*
- *if $\Delta'(1/2) < 0$, $\Delta(0) < 0$, both NAG and AG are geographical equilibria,*
- *if $\Delta'(1/2) > 0$, $\Delta(0) > 0$, only PAG is a geographical equilibrium.*

Proof. Using the expressions in (2.18), after some simplifications, one finds

$$\Delta(x) = \frac{(1-2x)((2a+4b)(x^2-x)+b)}{2(\sigma-\mu)(x+\tau^{\sigma-1}(1-x))(x\tau^{\sigma-1}+1-x)(x+\lambda(1-x))(x\lambda+1-x)}\,, \tag{3.2}$$

where

$$\begin{aligned} a &= (1-\tau^{\sigma-1})^2(\alpha N\sigma(1-\lambda) - \mu L(1+\lambda)^2 + 2\alpha N\mu\lambda) + 4\alpha N\tau^{\sigma-1}(1-\lambda)(\sigma-\mu)\,, \\ b &= \mu\lambda(1-\tau^{\sigma-1})^2(2L-\alpha N) - 2\alpha N\tau^{\sigma-1}(1-\lambda)(\sigma-\mu)\,. \end{aligned}$$

In the text we have shown that, due to the symmetry of our economy, the study of $\Delta'(1/2)$ and $\Delta(0)$ is enough to characterize the presence of NAG and/or AG equilibria.

The possibility of interior equilibria with $x \neq 1/2$ depends both on the zeros of $\Delta(x)$ and its derivative there.

We can restrict our analysis to the signs and derivatives of $N(x)$, the numerator of $\Delta(x)$. Indeed according to our hypothesis $\sigma > \mu$ and $\tau, \lambda \in (0, 1]$ so that the denominator of $\Delta(x)$ is always positive. This implies that both the signs and zeros of the rent difference are equal to the signs and zeros of its numerator, and that the sign of the rent difference derivative evaluated at its zeros is equal to the sign of the numerator rent difference derivative evaluated at the same points.

$N(x)$ is a cubic function symmetric around $x = 1/2$. Its parameterization in terms of a and b has been chosen so that

$$a = N'(1/2) \tag{3.3}$$

$$b = N(0)\,. \tag{3.4}$$

As a result when a and b are both positive, the derivative of the rent difference computed at $x = 1/2$ is positive, the rent difference at $x = 0$ is positive and, by symmetry, negative at $x = 1$. Given that $N(x)$ is a cubic polynomial, it must cross zero in two other points in the interval $[0, 1]$. Moreover at these two internal roots the marginal rent difference must be negative, so that these points are indeed geographical equilibria corresponding to PAG. It turns out that this is the only sign combination for which the economy is in a PAG status, as can be easily checked by repeating the same reasoning for all the other sign combinations of a and b (see also Fig. 2).

The asymmetric interior equilibria x^+ and x^- corresponding to PAG are the solutions of the second order equation $(2a + 4b)(x^2 - x) + b = 0$, that is,

$$x^{\pm} = \frac{1}{2}\left(1 \pm \sqrt{\frac{a}{a+2b}}\right)\,,$$

which, in fact, are in the interval $(0, 1)$ only when a and b have the same signs.

Finally notice that due to (3.3-3.4) conditions for the existence of PAG can be given in terms of $N(0)$, and thus $D(0)$, and $N'(1/2)$, and thus $D'(1/2)$, rather than a and b.

□

The previous proposition is silent on the limit cases when $\Delta'(1/2)$ and/or $\Delta(0)$ are equal to zero. We can analyze the 5 different cases (both coefficients equal to zero, one of the two equal to zero while the other is positive, the same when the other is negative) with the help of Fig. 2. In the special case when both $\Delta'(1/2)$ and $\Delta(0)$ are zero, rents are equal for any distribution of capital, that is, there is a continuum of interior equilibria. When $\Delta'(1/2)$ changes sign from negative to positive while $\Delta(0)$ is positive the economy transits from NAG to PAG. Otherwise, when $\Delta'(1/2)$ changes sign while $\Delta(0)$ is negative, it is only the existence of the NAG equilibria to be at stake, while AG is always stable and no other interior equilibrium appears.[5] Importantly, in the former case the transition between non agglomeration and

[5] When an explicit dynamic is added, using the language of bifurcation theory, the two phenomena are known respectively as a sub-critical and supercritical pitchfork bifurcation.

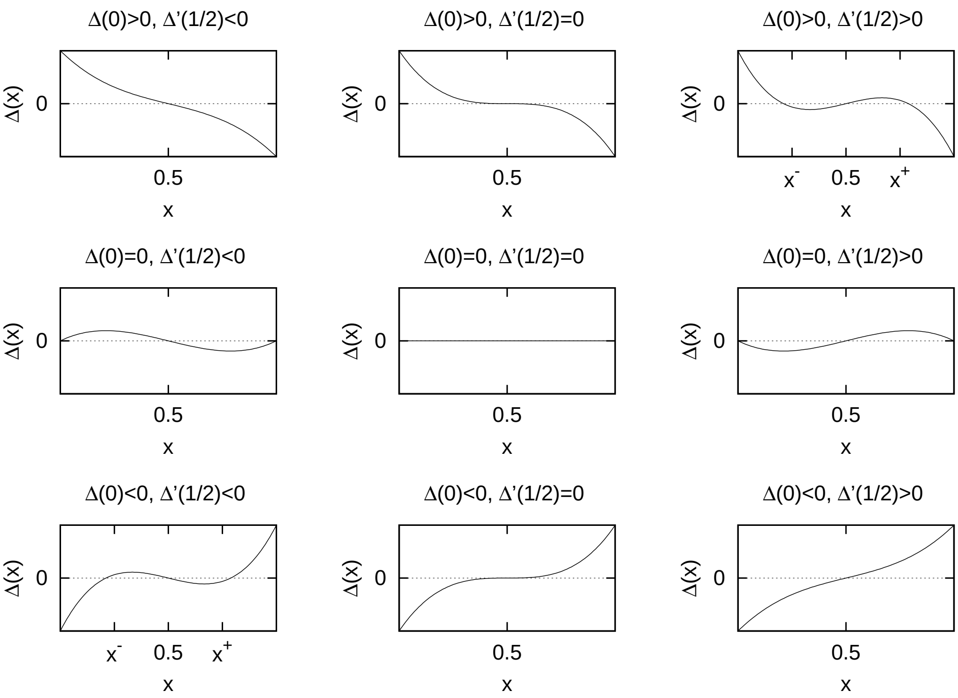

Figure 2: Capital rent differences for all the sign combinations of the coefficients $\Delta'(1/2)$ and $\Delta(0)$.

agglomeration is smooth and does not exhibit the typical hysteresis phenomenon associated with the latter. The same type of argument is valid when it is $\Delta(0)$ that changes its sign. The transition between full agglomeration and non full agglomeration is smooth when $\Delta'(1/2) > 0$ and exhibits hysteresis otherwise.

4 The effect of trade and technological openness

Having characterized the possible geographical equilibria of our economy, it remains to specify how their emergence depends on the parameters of the economy. More specifically we will be concerned with the influence of the inter-regional spillover and of the freeness of trade, taking consumer preferences, costs structure, number of firms and households as given and subject to the two restrictions (2.17) and (3.1). For this purpose, we translate the geographical equilibria conditions given above in Proposition 3.1 in terms of the more directly interpretable policy parameters λ and τ. This is the content of the following

Proposition 4.1. *Consider the rent difference* $\Delta(x)$ *as in* (3.2)*, and assume that* $\tilde{S} > 0$ *and* $S < min\{\tilde{S}, \bar{S}\}$*. There exist two functions* $\tau^a(\lambda)$ *and* $\tau^b(\lambda)$ *which map the interval* $(0, 1]$ *in the interval* $(0, 1]$ *such that*

$$\begin{aligned} \Delta'(1/2) \gtreqless 0 \quad &\Leftrightarrow \quad \tau \gtreqless \tau^a(\lambda) \\ \Delta(0) \gtreqless 0 \quad &\Leftrightarrow \quad \tau \lesseqgtr \tau^b(\lambda)\,, \end{aligned} \tag{4.1}$$

These two functions read

$$\tau^a(\lambda) = \left(1 + \Gamma(\lambda) - \sqrt{\Gamma^2(\lambda) + 2\Gamma(\lambda)}\right)^{\frac{1}{\sigma-1}} , \quad (4.2)$$

$$\tau^b(\lambda) = \left(1 + \Theta(\lambda) - \sqrt{\Theta^2(\lambda) + 2\Theta(\lambda)}\right)^{\frac{1}{\sigma-1}} , \quad (4.3)$$

with

$$\Gamma(\lambda) = \frac{\alpha N(1-\lambda)(\sigma-\mu)}{\mu\lambda(2L-\alpha N) + \frac{1-\lambda}{2}(\mu L(1-\lambda) - N\alpha\sigma)} ,$$

$$\Theta(\lambda) = \frac{\alpha N(1-\lambda)(\sigma-\mu)}{\mu\lambda(2L-\alpha N)} .$$

Moreover it holds $\tau^b(1) = \tau^a(1) = 1$, $\tau^b(0) = 0$ *and, when* $\lambda \gtreqless \tilde{\lambda} = 1 - \frac{N\alpha\sigma}{\mu L}$, $\tau^b(\lambda) \gtreqless \tau^a(\lambda)$.

Proof. As with the previous proof we characterize the signs of $\Delta(x)$ and $\Delta'(x)$ by looking at the signs of its numerator $N(x)$ and its derivative $N'(x)$. The formula for $N'(1/2)$ and $N(0)$ found in (3.3-3.4) are polynomial of second order in $\tau^{\sigma-1}$. The equations $\Delta'(1/2) = 0$ and $\Delta(0) = 0$ can thus be solved in terms of τ giving

$$\tau^a_\pm(\lambda) = \left(1 + \Gamma(\lambda) \pm \sqrt{\Gamma^2(\lambda) + 2\Gamma(\lambda)}\right)^{\frac{1}{\sigma-1}} ,$$

$$\tau^b_\pm(\lambda) = \left(1 + \Theta(\lambda) \pm \sqrt{\Theta^2(\lambda) + 2\Theta(\lambda)}\right)^{\frac{1}{\sigma-1}} ,$$

with

$$\Gamma(\lambda) = \frac{\alpha N(1-\lambda)(\sigma-\mu)}{\mu\lambda(2L-\alpha N) + \frac{1-\lambda}{2}(\mu L(1-\lambda) - N\alpha\sigma)} ,$$

$$\Theta(\lambda) = \frac{\alpha N(1-\lambda)(\sigma-\mu)}{\mu\lambda(2L-\alpha N)} .$$

Provided that $S = N/L < \bar{S}$, which is always the case given our assumptions, one can show that both $\Gamma(\lambda)$ and $\Theta(\lambda)$ are positive for any value of λ. This, in turn, implies that both τ^a_+ and τ^b_+ are larger than 1 for every λ in $[0,1]$, whereas τ^a_- and τ^b_- are two functions from $[0,1]$ to $[0,1]$ and corresponds to $\tau^a(\lambda)$ and $\tau^b(\lambda)$ in (4.2-4.3). Taking limits it is easy to show that $\tau^b(1) = \tau^a(1) = 1$ and $\tau^b(0) = 0$. Furthermore substituting for $\tilde{\lambda}$ in $\Theta(\lambda)$ and $\Gamma(\lambda)$ it is immediate to see that $\tau^a(\tilde{\lambda}) = \tau^b(\tilde{\lambda})$ and that $\tau^b(\tilde{\lambda}) \gtreqless \tau^a(\tilde{\lambda})$ when $\lambda \gtreqless \tilde{\lambda}$. Notice at last that since $S < \bar{S}$ it is $\tilde{\lambda} \in (0,1)$. □

The dependence of geographical equilibria on τ and λ is the consequence of their effect on $\Delta'(1/2)$ and $\Delta(0)$ which, in turn, depends on the trade-off between negative pecuniary and positive technological externalities. Figs. 3-4 bring together Propositions 3.1 and 4.1 and show for which values of the policy parameters λ and τ agglomerated, non-agglomerated, or partially agglomerated economies are observed. Despite in Fig. 3 the curves $\tau^a(\lambda)$ and $\tau^b(\lambda)$ are plotted for specific values of the economy parameters (L, N, μ, σ and α), their behavior and, in particular, the regions they identify are general properties of the model. Indeed it always holds that $\tau^b(0) = 0$, $\tau^b(1) = \tau^a(1) = 1$, $\tau^b(\tilde{\lambda}) \gtreqless \tau^a(\tilde{\lambda})$ when $\lambda \gtreqless \tilde{\lambda}$, and $\tilde{\lambda} \in (0,1)$. For the same reason the "bifurcation" phenomena illustrated in Fig. 4 are also general.

For high values of τ, $\tau > \tau^b(\lambda)$, the technological externality dominates, agglomeration on either sides is a geographical equilibrium and AG are observed. Conversely, for low values of

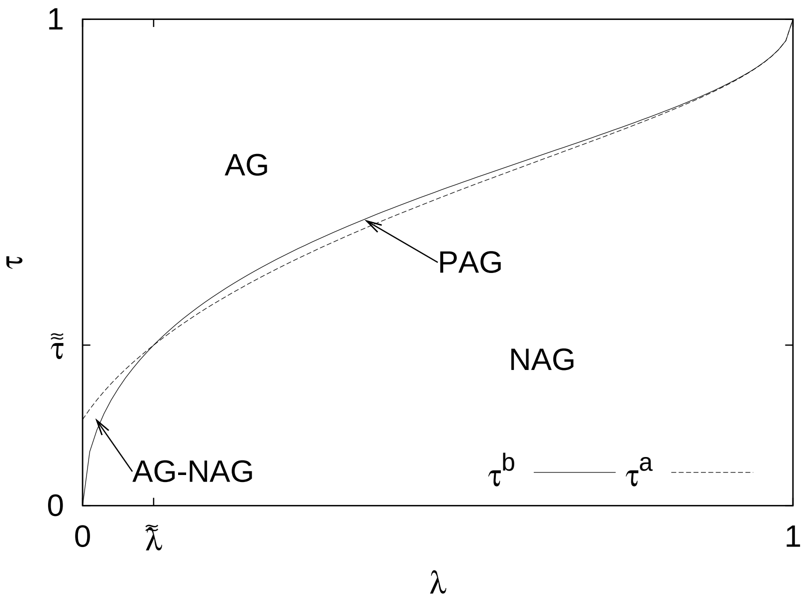

Figure 3: Existence of different type of geographical equilibria in the plane (λ, τ). The economy parameters are $L = 400$, $N = 150$, $\sigma = 3$, $\mu = 0.5$, $\alpha = 0.4$. The dotted and continuous curves are $\tau^a(\lambda)$ and $\tau^b(\lambda)$ respectively.

τ, $\tau < \tau^a(\lambda)$, the pecuniary externality dominates and the outcome is NAG. Firms, and thus capital, distribute equally between the two regions.

For intermediate values of transportation costs, geographical equilibria coexist and two different scenarios are possible, depending on the values of λ and τ. For low inter-regional spillovers, $\lambda < \tilde{\lambda}$, there exist values of the trade cost $\tau \in (\tau^b(\lambda), \tau^a(\lambda))$ such that the economy can be either in a NAG or AG configuration (c.f. the left panel of Fig. 4). When this is the case, the transition from NAG to AG, due to the opening-up of the economy, is abrupt. Moreover the economy shows hysteresis, that is, once the transition has occurred it is not the case that going back to a lower freeness of trade brings the economy back to its non-agglomerated state.

Things are different for higher inter-regional spillovers $\lambda > \tilde{\lambda}$, as shown in the right panel of Fig. 4. This time AG and NAG are still associated respectively with high and low values of τ, but the transition between the two equilibria is smoother. Indeed, for intermediate values of τ, between $\tau^a(\lambda)$ and $\tau^b(\lambda)$, two asymmetric geographical equilibria emerge, collapsing on the border (interior symmetric) equilibria as τ increases (decreases). These distributions represent PAG configurations: due to local spillover, the partial concentration of modern goods production is advantageous but, due to relatively high transportation costs, a further agglomeration is not beneficial as would only increase competition in the crowded location without enough profits coming from an increased demand in the other region.

5 Welfare Analysis

So far we have assumed that capital moves in order to maximize its rent, rather than households real income. This begs the question of what happens to household utilities, that is, to their welfare level. Household's utility in each location can be written as total income divided by

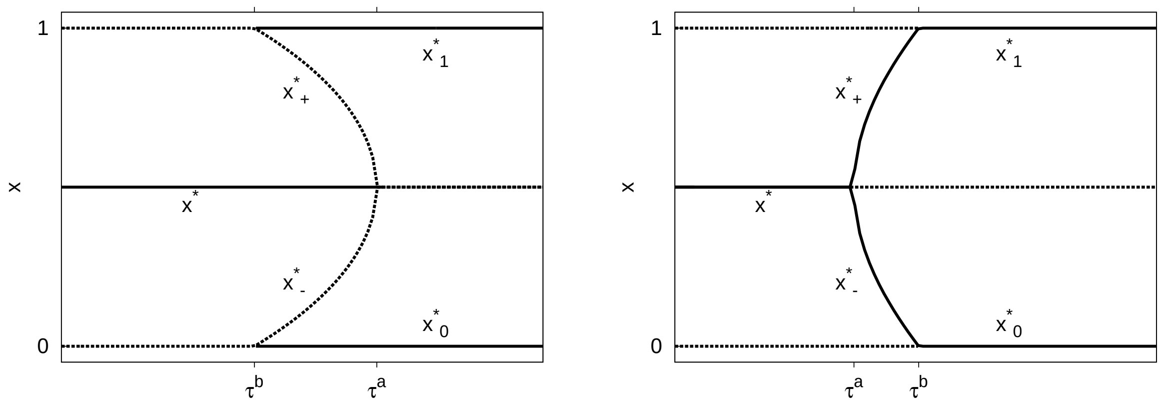

Figure 4: Geographical equilibria (solid lines) and non-equilibria equal capital rents or border distributions (dotted lines) as a function of τ. Left panel: $\lambda < \tilde{\lambda}$. Right panel: $\lambda > \tilde{\lambda}$.

the price index

$$\begin{cases} W_1(x) = \dfrac{I(x)}{P_1(x)} \\ \\ W_2(x) = \dfrac{I(x)}{P_2(x)}\,, \end{cases} \tag{5.1}$$

where $I(x)$ is given in (2.16-2.19),

$$P_1(x) = \left(\frac{\beta\sigma}{\sigma - 1} \left(Nx + N(1-x)\tau^{\sigma-1} \right)^{\frac{-1}{\sigma-1}} \right)^{\mu}\,, \tag{5.2}$$

and $P_2(x) = P_1(1-x)$.

The properties of $I(x)$ and $P_{1,2}(x)$ (c.f. the proof of Prop. 5.1) imply that the welfare difference

$$\Delta W(x) = I(x) \frac{P_2(x) - P_1(x)}{P_1(x)P_2(x)}$$

is negative in $x = 0$, zero only in $x = 1/2$ and positive in $x = 1$. As a result each household is better off if firms agglomerate in his/her own region. In this case local household does not pay transportation costs for modern goods and, due to the technology externality, their income benefits from the strongest possible abatement of fixed costs.

Whereas agglomeration is clearly beneficial for the region which happens to host the modern sector, it is not clear whether it is beneficial also for the whole economy. This is an important issue in a model like ours, where regions are ex-ante identical and where workers are not mobile. Finding an answer requires to investigate what happens to the welfare of the region that specializes in the traditional sector. On the one hand, households living there have to import all the modern goods so that, due to transportation costs, they have higher real prices. On the other hand, their nominal income is the same as that of households located in the modern region, and they also profits from higher capital rents. The overall result depends on the relative strength of these two effects, which in turn are related to both trade costs and inter-regional spillovers, that is, to market and technological openness.

Welfare analysis clearly depends on which type of welfare aggregating function one considers. Since we are mainly concerned with welfare levels in the traditional region, the Max-Min formulation seems the most appropriate. We define total welfare to be equal to the minimal welfare level between the two regions:

$$W_T(x) = \min\{W_1(x), W_2(x)\} \ . \tag{5.3}$$

Given that benefits of agglomeration spill also to the traditional region, it may well be the case that agglomeration is the best outcome, also under such an egalitarian definition of total welfare as (5.3). The following proposition rules when this is the case

Proposition 5.1. *Consider total welfare as in* (5.3)*. For any given value of the inter-regional spillover* λ*, provided that transportation costs* τ *are such that*

$$\tau > \tau^w(\lambda) = \left(2\left(\frac{L - \frac{N\alpha}{2}}{L - \frac{N\alpha}{1+\lambda}}\right)^{\frac{\sigma-1}{\mu}} - 1\right)^{-\frac{1}{\sigma-1}},$$

AG is the global welfare maximum. Otherwise, when $\tau < \tau^w(\lambda)$*, NAG is a global maximum. When* $\tau = \tau^w(\lambda)$ *both NAG and AG are global maxima.*

Proof. The proof relies on the properties of income $I(x)$ and of price indexes $P_{1,2}(x)$. For the income function defined in (2.16-2.19) one can easily show that $I(x) = I(1-x)$, $I'(1/2) = 0$, $I'(x) \gtreqless 0$ when $x \gtreqless 1/2$, and $I''(x) > 0$. For the price index functions in (5.2) it holds $P_1(x) = P_2(1-x)$, $P_1'(x) = -P_2'(x) < 0$, and $P_1''(x) = P_2''(x) > 0$. Using these properties we can re-write the expression of the Max-Min welfare as

$$W_T(x) = \begin{cases} W_1(x) & x \le \frac{1}{2} \\ W_2(x) & x \ge \frac{1}{2} \, . \end{cases}$$

Due to the symmetry of the economy, we can restrict our attention to the maxima of $W_2(x)$ in the interval $[1/2, 1]$. Given the behavior of $I(x)$ and $P_2(x)$ it holds both that $W_2'(1/2) < 0$ and that there exists at most one value of $x \in [1/2, 1]$ where $W_2'(x) = 0$. As a result the global maxima of the continuous and differentiable function $W_2(x)$ in the interval $[1/2, 1]$ are on its border, that is, either $x = 1/2$ or $x = 1$. In order to determine when one or the other prevail, we compare their welfare level and find

$$\frac{W_2(1)}{W_2(1/2)} \gtreqless 1 \Leftrightarrow \tau \gtreqless \tau^w(\lambda) = \left(2\left(\frac{L - \frac{N\alpha}{2}}{L - \frac{N\alpha}{1+\lambda}}\right)^{\frac{\sigma-1}{\mu}} - 1\right)^{-\frac{1}{\sigma-1}},$$

which proofs the proposition. □

AG is a welfare maximum provided that $\tau > \tau^w(\lambda)$. In this case the profits in the agglomerated region are so high that they offset the loss due to a high price index in the traditional region. High τs and low λs are in fact, respectively, lowering the price index difference between AG and NAG and increasing the gains in terms of capital rents due to agglomeration. Notice that since $\tau^w(\lambda)$ is an increasing function of λ, the minimal freeness of trade τ sufficient to make AG the welfare maximum is increasing with the strength of the inter-regional spillover λ.

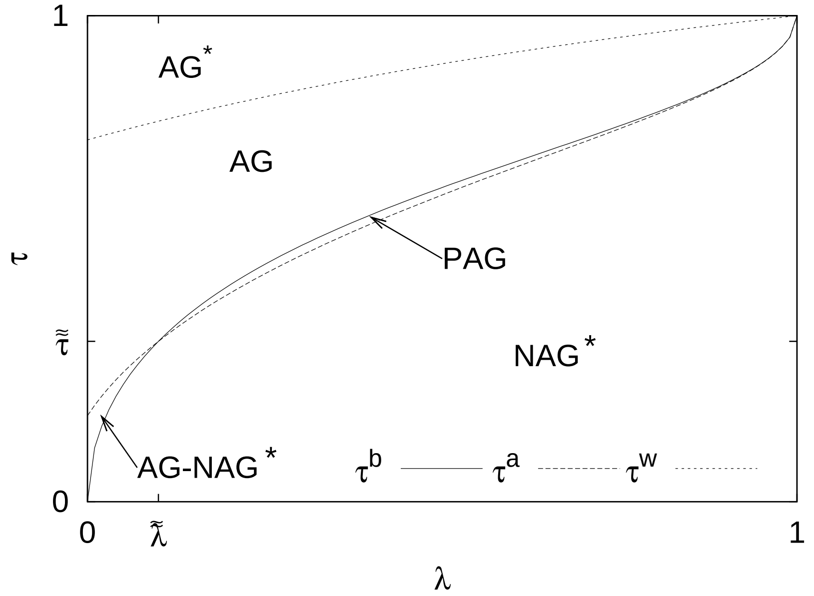

Figure 5: Geographical equilibria of Fig. 3 which are also total welfare maximizers are marked with a *. They are above the line $\tau^w(\lambda)$ in case of agglomerated economies and below it for non-agglomerated economies. The parameters are the same as in Fig. 3.

Conversely, when $\tau < \tau^w(\lambda)$ the welfare maximum is given by NAG. Given these differences in welfare levels, when does the geographical economic equilibrium arising from firm rents maximizing behavior lead to a total welfare maximum?

Figure 5 tries to answer this question putting together the results from Propositions 4.1 and 5.1. In the upper left area, above $\tau^w(\lambda)$, agglomerated economies are both a geographical equilibrium and the welfare maximum. In the lower right area, below $\tau^a(\lambda)$, the same is true for non-agglomerated economies. In Fig. 5 starred labels denote those geographical equilibria which are welfare maximizers. For any value of the inter-regional spillover there also exists an intermediate range of trade costs where NAG is the welfare optimum but the economic equilibria are AG or PAG. There seems to be no continuous path that links the upper-right area AG* with the lower-left area NAG* . The following lemma shows that this is a general result. By proving that the curve $\tau^w(\lambda)$ lies always above $\tau^a(\lambda)$, apart in the point $(\lambda = 1, \tau = 1)$ where they coincide, it shows that, unless $\lambda = 1$, it is never possible to move from NAG* to AG* following a path where the total welfare is always maximum.

Lemma 5.1. *Provided that $\tilde{S} > 0$ and $S < min\{\tilde{S}, \bar{S}\}$, it holds that*

$$\tau^a(\lambda) < \tau^w(\lambda)\,, \quad \textit{for every} \quad \lambda \in (0,1)\,, \tag{5.4}$$

and $\tau^a(\lambda) = \tau^w(\lambda)$ when $\lambda = 1$.

Proof. By evaluating $\tau^a(\lambda)$ and $\tau^w(\lambda)$ in $\lambda = 1$, one immediately sees that they are both equal to one. The rest of the statement has been proved numerically. We have defined a grid of 500 values of σ in $(1, 100]$, 100 values of μ in $(0, \sigma/(2\sigma - 1))$, which ensures that $\tilde{S} > 0$, 100 values of αS in $(0, \min\{\tilde{S}, \bar{S}\})$, and 500 values of λ in $[0, 1)$. We have checked that for every values of parameters and λ in the grid it holds that

$$\frac{d\tau^w(\lambda)}{d\lambda} < \frac{d\tau^a(\lambda)}{d\lambda} \qquad \text{when} \quad \lambda < \tilde{\lambda}$$

$$\frac{d\tau^w(\lambda)}{d\lambda} < \frac{d\tau^b(\lambda)}{d\lambda} \qquad \text{when} \quad \lambda \geq \tilde{\lambda}$$

which together with the fact that all these curves are equal to one when $\lambda = 1$, are monotonic, and $\tau^a(\lambda) \gtreqless \tau^b(\lambda)$ when $\lambda \gtreqless \tilde{\lambda}$, see Proposition 4.1, proves the result. □

5.1 Welfare enhancing policies

Having derived the total welfare for any value of "open-ness", the next concern is how an economy can move toward the line $\tau = 1$, that is, the locus of parameters where the total welfare is highest. Lemma 5.1 tells us that when the starting point is a low value of both freeness of trade and inter-regional knowledge spillover there is not a continuous path that brings the economy to the highest welfare without passing through an area where the geographical equilibrium is not a welfare maximum. In this case it is best to aim at reaching the full openness state where both λ and τ are both one, rather than any point on the line $\tau = 1$, even though all these states have the same welfare.

In general, under which conditions it is more beneficial to embrace policies that improve the freeness of trade and when, instead, it is better to go for inter-regional knowledge spillovers? To answer this question assume the government, by taxing households income I, can implement policies g which increase the level of freeness of trade τ and/or the inter-regional knowledge spillovers λ.

When the policy does not entail a change in the geographical equilibrium its effect on the total welfare for $x \leq 1/2$ can be computed as[6]

$$\frac{\partial W_T(x)}{\partial g} = \frac{\partial W_1(x)}{\partial g} = \frac{1}{P_1(x)} \frac{\partial I(x)}{\partial g} + \frac{\partial W_1(x)}{\partial \lambda} \frac{\partial \lambda}{\partial g} + \frac{\partial W_1(x)}{\partial \tau} \frac{\partial \tau}{\partial g} \tag{5.5}$$

where $\partial I(x)/\partial g$ stands for the cost of the policy. Under the assumption that a given amount of money spent by the policy g has the same impact on τ and λ, that is, $\partial \lambda / \partial g = \partial \tau / \partial g$, evaluating whether it is more welfare enhancing to increase τ or λ amounts to compare $\frac{\partial W_1(x)}{\partial \lambda}$ with $\frac{\partial W_1(x)}{\partial \tau}$, at the different geographical equilibria.[7] This is the result of the following lemma.

Lemma 5.2. *Consider the total welfare $W_T(x)$ as in* (5.3)*. Provided that the economy is in a AG state, it is always more beneficial to increase the freeness of trade τ rather than the inter-regional spillover λ. Otherwise, when the economy is in a NAG state, having defined*

$$\lambda^+(\tau) = \frac{\frac{N\alpha}{L} + \sqrt{\left(\frac{N\alpha}{L}\right)^2 + 4\frac{N\alpha}{L}\frac{1+\tau^{\sigma-1}}{\mu\tau^{\sigma-2}}}}{2} - 1\,, \tag{5.6}$$

[6]Given the symmetry of the welfare function around $x = 1/2$, one can easily compute the effect of the policy also for $x \geq 1/2$.

[7]The evaluation of the policy impact can be complicated by a variation in the geographical equilibrium distribution of the economy due to its effect on the integration λ and/or τ. This occurs at the non-generic direct transition between NAG and AG and in the generic case of PAG, when the degree of agglomeration depends continuously on the values of these parameters (see e.g. the left panel of Fig. 4). Since the parameters space where this dependence occurs is small, as can be seen in Fig. 3, we will skip it at this stage of the analysis.

it is more beneficial to increase the freeness of trade τ when $\lambda > \lambda^+(\tau)$, to increase the inter-regional spillover λ when when $\lambda < \lambda^+(\tau)$, and indifferent when $\lambda = \lambda^+(\tau)$.

Proof. The gradient of the total welfare at the two geographical equilibria corresponding to NAG and AG is:

$$\frac{\partial W_T(x)}{\partial \lambda} = \begin{cases} \left(\frac{N\alpha\sigma}{(\sigma-\mu)L(1+\lambda)^2}\right)\left(\frac{\sigma-1}{\beta\sigma}\left(\frac{N}{2}\right)^{\frac{1}{\sigma-1}}\right)^{\mu}(1+\tau^{\sigma-1})^{\frac{\mu}{\sigma-1}}\,, & x = 0.5\,, \quad \text{(NAG)} \\ 0 & x = 0,1\,, \quad \text{(AG)} \end{cases}$$

$$\frac{\partial W_T(x)}{\partial \tau} = \begin{cases} \left(\frac{\sigma\mu}{\sigma-\mu} - \frac{N\alpha\sigma\mu}{(\sigma-\mu)L(1+\lambda)}\right)\left(\frac{\sigma-1}{\beta\sigma}\left(\frac{N}{2}\right)^{\frac{1}{\sigma-1}}\right)^{\mu}(1+\tau^{\sigma-1})^{\frac{\mu}{\sigma-1}-1}\tau^{\sigma-2}, & x = 0.5\,, \quad \text{(NAG)} \\ \left(\frac{\sigma\mu}{\sigma-\mu} - \frac{N\alpha\sigma\mu}{(\sigma-\mu)2L}\right)\left(\frac{\sigma-1}{\beta\sigma}(N)^{\frac{1}{\sigma-1}}\right)^{\mu}\tau^{\mu-1}\,, & x = 0,1\,, \quad \text{(AG)} \end{cases}$$

Since for AG economies $\partial W_T/\partial\lambda$ is zero the best policy is always to increase the freeness of trade. The result for NAG economies follows from the comparison of the two components of the gradient evaluated there. The expression to be evaluated gives rise to a quadratic equation in λ whose only possibly positive root is $\lambda^+(\tau)$ as in (5.6). □

When the modern sector is agglomerated, fixed production costs abatement does not depend on the inter-regional spillovers, and neither do households welfare. As a result policies that increase λ have no effects and it is preferable to improve the freeness of trade τ. Conversely, when the modern sector is evenly spread between the two regions, Lemma 5.2 shows that it is more welfare improving to increase inter-regional knowledge spillovers when $\lambda < \lambda^+(\tau)$ and to increase freeness of trade τ otherwise. Fig. 6 summarizes these results and plots the gradient of the welfare function on the plane (λ, τ) for our benchmark choice of the economy parameters.

6 Conclusion

We have set up an analytical model with capital mobility, workers inter-sectoral mobility and inter-regional immobility, and where agglomeration is due to technological externalities. These externalities can be interpreted, for example, as shared costs for the creation of infrastructures or human capital. Due to the analytical solvability of our model, we have been able to compute the geographical equilibria and fully characterize their dependence on the trade-off between technological and pecuniary externalities, as regulated by transportation costs and inter-regional spillovers, and discuss their implications for total welfare.

Our analysis confirms previous findings by Baldwin and Forslid (2000) about the stabilizing nature of (knowledge) spillovers: the higher the spillover the larger the interval of transportation costs which lead to firms equidistribution. If the spillover is high enough, there exists a smooth equilibrium transition between agglomeration and equidistribution, with partly agglomerated economy for intermediate values of transportation costs. In this case an opening of inter-regional trade does not entail an abrupt reallocation of economic activities neither the hysteresis effect, typical of NEG model, which locks the economy in a core-periphery equilibrium also if higher trade costs are reintroduced.

Welfare analysis reveals that for a relatively large part of the (λ, τ) parameter space, even if the agglomerated outcome represents the geographic equilibrium, it generates less welfare in the periphery region than in the core. Since, in any case, the level of welfare of the periphery

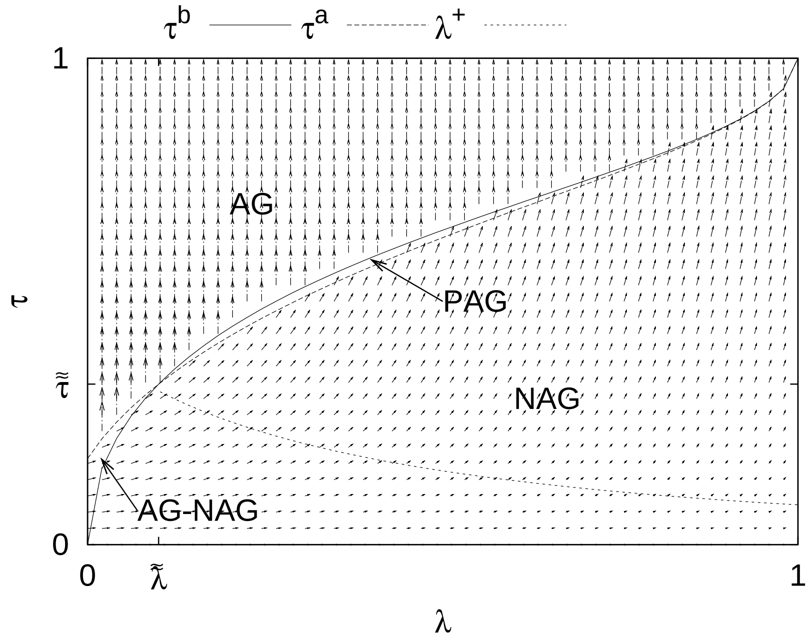

Figure 6: Gradients of the welfare function for NAG and AG economies. When the economy is in a AG state it is always better to increase freeness of trade τ. When the economy is in a NAG state it is better to increase the freeness of trade when τ is above the line $\lambda^{+}(\tau)$ and to increase inter-regional spillovers λ otherwise. The parameters are the same as in Figs. 3,5.

in the AG equilibrium increases with trade openness, for large enough level of τ this solution represent the welfare optimum for both regions. However, the existence of a large "welfare gap" makes the implementation of policies based on progressive opening of the economy difficult to implement.

On the other hand, the increase of the technological openness always improve the welfare level of both locations. When the level of knowledge sharing is low, its increase represent, from the point of view of the social planner, the best policy. Beside the positive effect on welfare levels, the increase of λ has also another advantage: for an economy with strong knowledge/technological integration, the "welfare gap" between agglomerated and non-agglomerated distribution is smaller and shallower and, consequently, policy geared toward markets integration are easier and less costly to implement.

In practice an increase in technological openness can be obtained by improving global means of information sharing, developing joint education programs, unifying norms and requirements affecting economic activities and the institutional constraints imposed to their markets. All these policies have the effect of improving global efficiency by avoiding replicated efforts and by abating knowledge barriers adding to transaction costs. Concluding, whereas freeness of trade leads, per-se, to sudden agglomeration, knowledge-based linkages favor a smoother transition between different levels of firms concentration and ultimately lead to a less uneven distribution of welfare.

References

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

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and R.-N. F. (2003). *Economic Geography and Public Policy.* Princeton University Press, Princeton.

Bottazzi, G., G. Dosi, G. Fagiolo, and A. Secchi (2008). Sectoral and geographical specificities in the spatial structure of economic activities. *Structural Change and Economic Dynamics 19*, 189 – 2002.

Brenner, T. (2006). Identification of local industrial clusters in Germany. *Regional Studies 40*, 991–1004.

Devereux, M., R. Griffith, and H. Simpson (2004). The geographic distribution of production activity in Britain. *Regional Science and Urban Economics 34*, 533–564.

Dixit, A. and J. Stiglitz (1977). Monopolistic competition and optimun product diveristy. *American Economic Review 67*(3), 297–308.

Dupont, V. and P. Martin (2006). Subsidies to poor regions and inequalities: some unpleasant arithmetic. *Journal of Economic Geography 6*, 223–240.

Ellison, G. and E. Glaeser (1997). Geographical concentration in U.S. manufacturing industries: a dartboard approach. *Journal of Political Economy 105*, 889–927.

Forslid, R. and G. Ottaviano (2003). An analytically solvable core-periphery model. *Journal of Economic Geography 3*, 229–240.

Grossman, G. and E. Helpman (1980). *Innovation and Growth in the World Economy.* MIT Press, Cambridge Mass.

Krugman, P. (1991a). *Geography and Trade.* The MIT Press, UK.

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

Marshall, A. (1920). *Principles of Economics.* London: Macmillan and Co.

Martin, P. (1999). Public policies, regional inequalities and growth. *Journal of Public Economics 73*, 85–105.

Martin, P. and G. Ottaviano (1999). Growing locations: Industry locations in a model of endogenous growth. *European Economic Review 43*, 281–302.

Martin, P. and C. A. Rogers (1995). Industrial location and public policy. *Journal of International Economics 39*, 335–351.

Maurel, F. and B. Sedillot (1999). A measure of the geographic concentration in french manufacturing industries. *Regional Science and Urban Economics 29*, 575–604.