Transport liberalization and regional imbalances with endogenous freight rates

Francesco Di Comite*

August 1, 2013

This paper proposes an analytically solvable New Economic Geography model to investigate the long-run industry reallocation patterns triggered by transport liberalization once the determination of freight rates has been endogenized. Two policy scenarios are considered: one where the regulator increases competition by increasing the number of carriers but imposes a unique tariff per route and one of complete deregulation which leaves carriers free to price discriminate. Under the complete deregulation scenario, profit maximizing carriers are shown to charge different prices in the two directions of the same route. Carriers are indeed shown to extract higher markups when delivering goods to the less competitive regions. This pricing behavior in the transport sector counterbalances standard agglomeration forces associated with lowering trade costs and can even generate dispersion if the number of firms or the substitutability between varieties are high enough.

JEL codes: R12 - R58 - O18.

Keywords: Transport liberalization; endogenous transport costs; regional imbalances; Home Market Effect.

^{*}Department of Economics (IRES), Université Catholique de Louvain, Joint Research Centre, European Commission. Email: francesco.di-comite@ec.europa.eu. The author is solely responsible for the content of the paper. The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

1 Introduction

In the absence of product differentiation and economies of scale, trade costs can be seen simply as a source of price distortion leading to a homogeneous loss of welfare for all the market participants. This consideration has often led policy makers to develop an attitude towards transport costs which can be epitomized as: the cheaper, the better. Lower transport costs are expected to yield a better integrated market which should result in improvements in citizens' welfare through the reduction in prices associated with a more intense competition in the final products' market. For example, the Lisbon Agenda clearly pushed decisively in this direction and openly celebrated the early success of transport liberalizations policies: "The first real advance in common transport policy brought a significant drop in consumer prices, combined with a higher quality of service and a wider range of choices" (European Commission, 2001).

But are these early welfare improvements meant to last in the long run? Do they depend on the market structure in the transport sector? And are the resulting gains evenly distributed across regions? This paper tries to answer these questions in the context of a New Economic Geography (NEG) framework in which products are differentiated and firms' interactions are described as in the monopolistic-competition tradition (Dixit and Stiglitz, 1977; Krugman, 1991) with the additional feature of an endogenous transport sector à la Behrens Gaigné and Thisse (2009) where firms are also allowed to price discriminate by direction of shipment. NEG models are typically adopted for yielding predictions on the spatial distribution of economic activities, which then map into location-specific market outcomes. The key underlying idea is that once firms are allowed to relocate from one region to the other to maximize their profits, then market outcomes are affected in both origin and destination markets because of an alteration in the competitive interactions of firms (Fujita, 1988; Ottaviano, Tabuchi and Thisse, 2002).

This paper investigates the spatial agglomeration patterns and welfare implications of a cost-reducing liberalization of the transport sector, with and without the possibility for carriers to price discriminate by the direction of shipment. Following the analysis of Behrens and Picard (2011), transport costs are not treated as parameters after liberalization, but they are determined endogenously by profit maximizing carriers providing an undifferentiated transport service. However, in this paper no constraints are imposed on the supply side of the transport sector and the focus is on the welfare impact of different transport liberalization policy options. Two scenarios are analyzed: the first is one of a regulated liberalization in which carriers have to be fix freight rates per route, independently of the direction of the shipment, which follows more closely the modeling strategy of Behrens Gaigné and Thisse (2009). The second is a complete deregulation, allowing for price discrimination based on the direction of the shipment closer to the modeling appproach of Behrens and Picard (2011) with the main difference that carriers are not forced to commit to capacities and offer transport services in both shipping directions.

Whereas the transport service is assumed to be undifferentiated, leading to a standard oligopolistic outcome à la Cournot (1838), the same is not true of the manufactured goods transported by the carriers. These are supplied in horizontally differentiated varieties produced by single-product identical firms operating in a monopolistic competition framework characterized by increasing returns to scale and a limited amount of market power granted from consumers' love for variety. In particular, the use of a variableelasticity-of-substitution utility functions (Ottaviano, Tabuchi and Thisse, 2002; Belleflamme, Picard and Thisse, 2000), allows to capture firms price discrimination across markets (as opposed to constant-elasticity-of-substitution preferences yielding constant markups on costs) and to conveniently treat transport costs as linear rather than iceberg. Linear transport costs have the notable advantage of allowing economists to explicitly model a competitive transport sector and study the impact of transport policies on the spatial distribution of economic activities and ultimately on welfare. One way to interpret the dynamics of a model with profit maximing carriers and manufacturers is to fram it in terms of the well-known Industrial Organization concept of double marginalization (Tirole, 1988), where the profits that the two types of agents can extract from the final consumers depend on the toughness of competition in the market where the final goods are eventually sold.

It is worth noting that in the framework explored in this paper there is no long-run reversal of the short-run gains in welfare, which instead is instead a common result in the NEG literature. What is even more surprising is that in the complete deregulation scenario transport liberalization can alleviate rather than exacerbate regional disparities. The intuition behind this latter result is that, as compared to the unique-bilateral-tariff setting, the tougher (softer) degree of competition in the core (peripheral) regions drives carriers to charge a lower (higher) price to firms supplying those market from the periphery (core), but this makes more convenient for firms to move to the periphery and enjoy softer competition in the core region and lower transport costs to serve the other.

The balanced geographical dstribution of economic ctivities resulting from the decentralized solution contrasts starkly with most results in the NEG literature, where externalities and spillovers often call for regulatory intervention to address market failures. In fact, a general NEG result is that any symmetric reduction in transport costs is bound to foster economic agglomeration of firms in the most affluent region, in the context of increasing returns to scale. This process can push the relative welfare of the different types in opposite directions in the long run, once firms' relocation is taken into account. In Behrens Gaigné and Thisse (2009), for example, it is stated that "the short-run benefits of deregulation could well be offset by the long-run costs triggered by making the distribution of economic activity less efficient." However, this concern is shown to be less justified in the case analyzed in this paper, i.e. when direction-specific transport price discrimination is taken into account.¹ Notice that by focusing on the demand for transport services and abstracting form a specific supply structure, the modeling approach of this paper can be straightforwardly applied to any sources of trade costs where prices are determined by pure profit maximizing agents, be them port or airport authorities, trade insurers or any other endogenously determined variable source of trade costs which can be destination- or direction-specific. It may even be applied to governments trying to maximize revenues derived from import tariffs, following the traditional Brander and Spencer (1984) intuition that optimal tariffs depend on the elasticity of the import demand, which in this paper's framework is variable and depends on the amount of industrial concentration.

The main policy message stemming from this paper is that different liberalization policies may be preferred depending on policy-makers' priorities. Each of the three analyzed regimes responds to a specific social need: a *unique-tariff liberalization* is shown to lower transport costs, especially from the periphery to the core and keep them low even from the core to the periphery, leading to the maximum welfare gains in the peripheral region but also to high levels of deindustrialization; a *complete deregulation* would grant relatively lower welfare gains to the poorest region and greater welfare gains to the core

¹It is worth stressing that in the model with capital mobility and no migration analyzed here there are no income effects. These would play in favor of agglomeration, but they would not alter the results substantially.

region, but at the same time the relatively lower transport costs from the periphery to the core would put some restraint on agglomeration; finally, a *high-cost regulated transport sector* will keep the spatial distribution of employment and production as dispersed as possible but would yield the worst possible outcome in terms of welfare to consumers in both type of regions.

The paper is organized as follows. In section 2, the model is developed in the short run and commented. In section 3 the relocation of firms in the long run is described. In section 4 the transport sector yielding endogenous freight rates is introduced in the model, under two different regulatory settings: complete deregulation and unique-bilateral-tariff liberalization. The two different regulatory regimes are compared in terms of welfare and geographic distribution of industrial activity in section 5. Finally, section 6 concludes.

2 The model

Consider an economy composed of two regions: a more populated core, H, and a less populated periphery, F (for notational purposes, the two regions are called i and j when considered in general terms). Consumers exhibit quadratic preferences and the economy is characterized by two sectors, one displaying constant returns to scale and using one unit of labor to produce the numéraire, the other exhibiting increasing returns to scale and using inter-regionally mobile capital and local labor to produce varieties of a differentiated good. When the differentiated good is produced in i and sold in j, it incurs a linear transport cost $t_{i,j}$, whereas $t_{i,i} = 0$. The value of $t_{i,j}$ is initially taken as a parameter (under a regulated transport sector) and independent of the direction of the shipment $(t_{i,j} = t_{j,i} = t)$; then, it is determined endogenously in the transport sector following two types of liberalization, one imposing $t_{i,j} = t_{j,i}$, which is called *unique-bilateral tariff*, and the other allowing carriers to price discriminate by the direction of shipment, $t_{i,j} \neq t_{j,i}$, which is called *complete deregulation*.

Here follows a description of the economy in the short run.

2.1 Consumers

The economy is inhabited by M identical consumers which are exogenously distributed in the two regions and are endowed with one unit of labor, L, which is geographically immobile, and one unit of capital, K, which can be invested in either region. The share of people living in region i is expressed by θ_i , i.e. $\theta_H \in [\frac{1}{2}; 1]$ and $\theta_F = 1 - \theta_H$, so that the absolute number of consumers living in that region is $M_i = \theta_i M$. To simplify the notation, in the rest of the paper whenever θ is left without subscript, it refers by convention to θ_H , which can be interpreted as a measure of demand concentration in the economy. The preferences of each consumer are captured by a standard quadratic utility function (Ottaviano *et al.*, 2002; Belleflamme *et al.*, 2000):

$$U_{i}(s) = \alpha \int_{s \in N_{i}} q_{s,i} ds - \frac{\beta}{2} \int_{s \in N_{i}} q_{s,i}^{2} ds - \frac{\gamma}{2} \left[\int_{s \in N_{i}} q_{s,i} ds \right]^{2} + q_{0}, \tag{1}$$

Where N_i is the mass of varieties present in region *i*, each variety being of negligible size for the market, $q_{s,i}$ is the amount of variety $s \in (0; N]$ consumed by each consumer *i*, the parameter α defines the intensity of preference accorded to the consumption of the differentiated good, as compared to the homogeneous one, q_0 , whose marginal utility is normalized to unity and used as the numéraire of the economy. The parameters $\beta \in$ $(0; \infty)$ and $\gamma \in [0; \infty]$ determine consumer's love for variety by capturing perceived horizontal differentiation of varieties (β) and the degree of substitutability (γ) between varieties.² Consumer are subject to the following budget constraint:

$$\int_{s \in N_{ii}} p_{s,ii} q_{s,ii} ds + \int_{s \in N_{ji}} p_{s,ji} q_{s,ji} ds + q_0 = y_i + \bar{q_0},$$
(2)

where $p_{s,ii}$ and $q_{s,ii}$ are the price (in terms of the numéraire) and quantities sold of a variety s of differentiated good bought by a consumer living in the same region as the producing firm; $p_{s,ji}$ and $q_{s,ji}$ are the price and quantities sold of a variety of differentiated good bought by a consumer living in a region different from the one where the producing firm operates; the parameter \bar{q}_0 represents the consumer's *i* initial endowment of homogeneous good (assumed to be large enough to allow the consumer to enjoy any level of consumption of the differentiated good); finally, y_i is consumers' nominal income earned through the provision of factors L in region *i* and K in either one of the two regions.³

²For a detailed discussion of the quadratic utility's parameters, see Di Comite *et al.* (2011)

³Notice that since wages are determined in the numéraire producing sector and profits are redistributed to consumers in the two regions, no differences arise between the two regions in terms of nominal income. The numéraire is assumed here to be freely traded, as common in NEG models, even if this assumption has been shown from Picard and Zeng (2005) to have stronger implications than generally thought. Indeed, the presence of transport costs in the homogeneous good's market (which is assumed to be an

Optimizing (1) subject to (2) with respect to $q_{s,ii}$ leads to the following linear demand function:

$$q_{s,ji} = \frac{\alpha - p_{s,ji} - \gamma Q_i}{\beta},$$

where $Q_i = \int_{s \in N_i} q_{s,i} ds$. This can be rewritten as

$$q_{s,ji} = \frac{\alpha\beta + \gamma N_i \bar{p}_i}{\beta(\beta + \gamma N_i)} - \frac{p_{s,ji}}{\beta},\tag{3}$$

where

$$\bar{p}_i = \frac{\int_{s \in N_i} p_{s,i} ds}{N_i}$$

is a price index capturing the average price of all the varieties of the differentiated good sold in region i.⁴

The linear demand function (3) encloses the idea that the demand of a certain variety s in market i falls when its price rises not only in absolute terms (*own price effect*, $p_{s,ji}$) but also with respect to the average price (*differential price effect*, P_i). However, these effects exhibit a different magnitude, as the own price effect $\left[-\frac{1}{\beta}\right]$ is always stronger than each cross price effect (which is negligible by definition) and than the sum of all the cross price effects $\left[\frac{\gamma N}{\beta(\beta+\gamma N)}\right]$. In addition, notice that this demand structure allows for variable elasticity of substitution in own and market prices:

$$\epsilon_{s,ji} = \frac{\delta q_{s,ji}}{\delta p_{s,ji}} \frac{p_{s,ji}}{q_{s,ji}} = \frac{p_{s,ji}}{\alpha\beta - p_{s,ji}(\beta + \gamma N_i) + \gamma N_i \bar{p}_i}.$$

2.2 Firms

Turning to the production side of the economy, only two factors of production are used in the production processes: a regionally mobile one, capital (K), and an immobile one, labor (L). The perfectly competitive constant-returns-to-scale sector produces the homogeneous good employing only labor. The monopolistically competitive differentiated manufacturing sector, with single-product firms operating under increasing returns to scale, employs both factors, capital in fixed amounts and labor proportionally to produc-

agricultural good in their case) turn out to be a rather important dispersion force.

⁴It is worth noting already at this point that in this model there is no heterogeneity in consumer demand or productive efficiency. Still, there will be two different prices in the market, as a share of varieties in i will be shipped from j and the transport costs will be partly passed thorough to final consumers.

tion. In the manufacturing sector, firms' profit can be expressed as follows:

$$\Pi_{s,i} = (p_{s,ii} - c)q_{s,ii}M_i + (p_{s,ij} - c - t)q_{s,ij}M_j - r_i f,$$
(4)

where c is the amount of labor needed to produce one unit of the differentiated good, t is the linear transport cost, taken as exogenous, r is the return on capital invested, and f is the amount of capital needed to set up a firm, which can be interpreted as a fixed entry cost. Plugging (3) into (4) the profit function can be rewritten as:

$$\Pi_{s,i} = (p_{s,ii} - c) \left[\frac{\alpha\beta - (\beta + \gamma N_i)p_{s,ii} + \gamma N_i \bar{p}_i}{\beta(\beta + \gamma N_i)} \right] M_i + (p_{s,ij} - c - t) \left[\frac{\alpha\beta - (\beta + \gamma N_j)p_{s,ij} + \gamma N_j \bar{p}_j}{\beta(\beta + \gamma N_j)} \right] M_j - r_i f.$$
(5)

Notice that the total number of firms in the economy, $N = N_i + N_j$, is a function of the amount of capital in the economy and the fixed entry cost, N = K/f, but it is split between the two regions according to the fraction of capital, λ , allocated to each region, so that:

$$N_i = \frac{\lambda_i K}{f} \quad ; \quad N_j = N - N_i. \tag{6}$$

For notational convenience, when there is no subscript λ refers to the core region, i.e. $\lambda = \lambda_H$, and it can be considered a measure of producers' agglomeration.

2.3 Market outcomes in the short run

As manufacturing firms compete in a monopolistic competition framework, they are assumed to maximize their profits, as captured by (5), through the choice of optimal prices in their destination markets: the domestic and the foreign. Given that regions are not fully integrated (in the sense that $t_{ij} > 0$), two different prices will emerge for the same product in the two markets, whose level of competition is captured by price indices.⁵ Considering that there is no other source of heterogeneity except geographic location, four segments can be identified in the market, as in each of the two regions there will be one type of firm facing transport costs and the other not. The segments are then HHand HF for the irms located in the core and FF and FH for the firms located in the

⁵Firms take price indices as exogenous because that each firm is assumed to be of a negligible size.

periphery.

The profit-maximizing prices chosen by the firms in region i are then:

$$p_{s,ii} = \frac{\alpha\beta + \gamma N_i \bar{p}_i}{2(\beta + \gamma N_i)} + \frac{c}{2} \quad ; \quad p_{s,ji} = p_{s,ii} + \frac{t}{2} = \frac{\alpha\beta + \gamma N_i \bar{p}_i}{2(\beta + \gamma N_i)} + \frac{c}{2} + \frac{t}{2}, \tag{7}$$

which is the standard outcome of a monopolist facing a linear demand function, the profit-maximizing price being just half of the highest possible price consistent with positive consumption levels, plus half of any marginal cost associated with production, c and shipment, t, if applicable.

As expected, the differentiated varieties' prices rise as consumers' bias toward the consumption of differentiated goods (α), marginal costs (c) or price indices (\bar{p}) increase. Product differentiation as well, as captured by the parameter β (the higher, the more variety in consumption is appreciated) plays an important role, as can be better understood by developing equation (7). Indeed, taking into account that

$$\bar{p}_i = P_i/N_i = \int_{s \in N_i} \frac{p_{s,i}}{N_i} ds = \frac{N_{ii}}{N_i} p_{s,ii} + \frac{N_{ji}}{N_i} p_{s,ji} = p_{s,ii} + \frac{N_{ji}}{N_i} \frac{t}{2} = p_{s,ii} + \lambda_j \frac{t}{2}$$
(8)

and expressing p^* only in terms of structural parameters (taking λ as a parameter too, at least in the short run), it can be seen that:

$$p_{s,ii}^{*}(t) = \frac{\beta(\alpha+c) + \gamma N_i(c+\lambda_j \frac{t}{2})}{2\beta + \gamma N_i} \quad ; \quad p_{s,ji}^{*}(t) = p_{s,ii}^{*} + \frac{t}{2}.$$
(9)

From (9) it can be noticed that as $\beta \to 0$ (or similarly as $\gamma \to \infty$) consumers' love for variety disappears and $p_{s,ii}^* \to c + \frac{\lambda_j}{2}t$ and $p_{ji}^* \to c + \frac{\lambda_j}{2}t + \frac{t}{2}$. This is exactly equal to the marginal cost of production plus a markup component deriving from the acknowledgment that a share λ_j of firms in the market is characterized by higher marginal costs of production and delivery, c + t, thus affecting the price index and relaxing price competition.

Similarly, average prices in market i can be expressed in terms of the structural parameters and transport costs:

$$\bar{p}_i = \frac{P_i}{N_i} = \frac{\alpha\beta + (c + \lambda_j t)(\beta + \gamma N_i)}{2\beta + \gamma N_i},\tag{10}$$

which confirms that, as $\beta \to 0$ (or $\gamma \to \infty$), then $\bar{p}_i \to c + \lambda_j t$. This result is explained

by the fact that a share λ_i of the varieties found in the region has marginal costs equal to c and a share λ_j of varieties has marginal costs equal to c + t. Remembering that that higher transport costs generate a less than proportional increase in final prices, it can safely be claimed that in the short run (i.e., when the spatial distribution of economic activities is considered fixed) high transport costs further aggravate the market distortions generated by the existence of increasing returns and damage consumers by reducing both consumers' surplus and firms' profits, which is a standard result in the NEG literature. In particular,

$$p_{s,ii}^*(t) = \bar{p}_i - \frac{t}{2}\lambda_j$$
; $p_{ji}^*(t) = \bar{p}_i + (1 - \lambda_j) = \bar{p}_i + \frac{t}{2}\lambda_i$,

from which two considerations derive. The first is that price differentials in price between the domestically produced and the imported varieties are directly related to the magnitude of transport costs. The second is that transport costs affect asymmetrically the optimal pricing of the two varieties. Since $\lambda_i > \lambda_j$, importers' deviation from the average price in region *i* (charging a higher price than the average) is higher than domestic firms' (charging a lower price than the average). In addition, from equation (7) it can be observed that prices of individual varieties and regional price indices can be seen as strategic complements, as shown in Figure 1.

Equilibrium prices, as expressed in (9), could also be plugged into the demand function (3), so as to obtain the equilibrium quantities only as a function of structural parameters:

$$q_{s,ii}^*(t) = \frac{\beta(\alpha - c) + \gamma N_i \lambda_j \frac{t}{2}}{\beta(2\beta + \gamma N_i)} \quad ; \quad q_{s,ji}^*(t) = \frac{\beta(\alpha - c) - \frac{t}{2}(2\beta + \gamma N_i \lambda_i)}{\beta(2\beta + \gamma N_i)}. \tag{11}$$

As expected, the transport cost, t, enters positively in q_{ii}^* and negatively in q_{ji}^* , but again asymmetrically. Indeed,

$$\frac{\partial q_{ii}^*}{\partial t} = \frac{\gamma N_i \frac{\lambda_j}{2}}{\beta(2\beta + \gamma N_i)} \quad ; \quad \frac{\partial q_{ji}^*}{\partial t} = -\frac{\beta + \gamma N_i \frac{\lambda_i}{2}}{\beta(2\beta + \gamma N_i)}. \tag{12}$$

This means that even if transport costs shift demand towards domestically produced goods at the expense of imported ones, they create less demand on the domestic segment than they destroy on the imported one. This implies that the total amount of consumption in the differentiated sector is reduced and this implies that prices in each region rise for both segments, as shown in equations (9) and (10).

As for the effects of industrial agglomeration, it can be noted that as long as $\lambda > \frac{1}{2}$, both prices and quantities are always lower in the bigger region than in the smaller one, when t is equal in the two directions. Therefore, tougher competition in the bigger region induced by lower transport costs benefits consumers in H by raising their real wage (because the nominal wage is normalized to unity, but goods' prices are declining). However, the same is not true for the firms located in H. Indeed, in the short run the profits of the firms located in the periphery increase in their export segment while the profits and market share of the domestic firms decrease in the local market.

2.4 Consumers' surplus

Based on the assumed preference structure shown in (3), the profit-maximizing pricing shown in (9) and the corresponding quantities sold in equation (11), it is possible to derive an expression to capture consumers' surplus, which is the conventional tool used to assess how the market structure is affecting consumers' welfare. The idea behind the analysis of the consumer surplus is that consumers in the economy would still buy some units of the differentiated good even if prices were higher, given that the marginal utility of consumption decreases in quantities. This means that at the actual equilibrium price, the inframarginal units consumed are paid less than the what the consumer would accept to paid for them, thus providing more utility than is being paid for.

Analytically, the utility associated with these gains can be quantified through the analysis of the indirect utility function derived from (1). It can be expressed in function of prices and income:

$$S_i = \frac{\alpha^2 N_i}{2(\beta + \gamma N_i)} - \frac{\alpha}{\beta + \gamma N_i} \int_{s \in N_i} p_{s,i} ds + \frac{\int_{s \in N_i} p_{s,i}^2 ds}{2\beta} - \frac{\gamma}{2\beta(\beta + \gamma N_i)} \left[\int_{s \in N_i} p_{s,i} ds \right]^2 + Y_i + \overline{q}_0.$$
(13)

The exact amount of consumer surplus associated with a particular market structure can then be obtained by just plugging the equilibrium prices and income in the equation (13), but the main interest lies not on in its value per se, but rather on its changes in response to policy or technology changes. It can be noticed in (13) that consumers' surplus is always negatively affected by price increases and the effect is stronger the higher are the quantities sold:

$$\frac{\partial S_i}{\partial p_{s,i}} = -\frac{\alpha\beta - (\beta + \gamma N_i)p_{s,i} + \gamma N_i \bar{p}_i}{\beta(\beta + \gamma N_i)} < 0.$$

Looking at (3), it is indeed worth remarking that

$$\frac{\partial S_i}{\partial p_{s,i}} = -q_{s,i}.$$

This implies that surplus is expected to be higher in the core region and every change in transport costs (t) or industrial agglomeration (λ) affecting prices maps directly into changes in the surplus of the consumers in the two regions.

2.5 Factors' remuneration

As mention earlier in the text, of the two factors of the model labor can be used to produce the two types of good, the homogeneous under constant returns to scale and the differentiated under increasing returns to scale, its supply being perfectly elastic at the wage level corresponding to the value of the homogeneous good (the numéraire). Thus, the resulting wage in nominal terms will be equal in the two sectors of the two regions.

As for capital, it is taken as fixed in the short run but it becomes mobile in the long run. This means that, after a shock in transport costs, remuneration can temporarily differ, but will eventually equalize across the two regions. Its remuneration is directly related to the operating profits generated by firms in the two regions. Indeed, as capital is the scarce resource in this economy, all the operating profits (the difference between revenues and salaries) are absorbed by its remuneration. This can be interpreted as the result of a bidding process in which any new entrant firm, if incumbents are making profits, has room to offer a slightly higher remuneration and so attract all the capital of the economy, thus leading to a fierce competition between firms to the advantage of capital holders. As a consequence, as far as there is free entry of enterprises in the heterogeneous goods' manufacturing market and no heterogeneity across firms, consolidated profits are equal to zero and the remuneration of capital in the two regions equate the operating profits:

$$r_{s,i} = \frac{(p_{s,ii} - m)q_{s,ii}M_i + (p_{s,ij} - m - t)q_{s,ij}M_j}{f},$$

which can be rewritten as:

$$r_{s,i} = \frac{M}{f\beta} \left[(p_{s,ii} - c)^2 \theta_i + (p_{s,ij} - c - t)^2 \theta_j \right].$$

From this expression, it can be understood where the trade-offs concerning capital remuneration stem from. Each variety is in fact extracting profits from two segments, the domestic and the foreign, each having a different number of consumers ($\theta_i M$ and $\theta_j M$), different marginal costs of production and delivery (c and c + t) and different local price indices yielding different prices ($p_{s,ii}$ and $p_{s,ij}$). This explains how, even in the absence of any technological difference, subsidy or barrier to trade, the two regions can reach very different levels of industrialization, just on the basis of differences in the consumption levels.

3 The long run: industrial agglomeration and regional imbalances

In the long run, capital is free to move between the two regions. Thus, it will flow from one region to the other until capital holders are indifferent between investing in one region or in the other, i.e. when $r_i = r_j$, which can be rewritten as

$$\left[(p_{s,ii}^*-c)^2\theta_i + (p_{s,ij}^*-c-t)^2(1-\theta_i)\right] = \left[(p_{s,jj}^*-c)^2(1-\theta_i) + (p_{s,ji}^*-c-t)^2\theta_i\right].$$

As long as transport costs t are taken as exogenous, it must be noted that this relation holds only if transport costs are not excessive, i.e. if there is trade between the regions. The maximum value consistent with the existence of international trade in the two directions, which can be called t^*_{trade} , can be computed as the value which ensures q_{ji} in (11) or, equivalently, p_{ji} in (9) to be positive. In terms of the structural parameters, it is

$$t_{trade} = \frac{\beta(\alpha - c)}{\beta + \frac{\gamma N}{2}}.$$
(14)

Focusing then on cases in which $t \leq t_{trade}$ and solving the capital equalization equation, it can be seen that following relation holds between the agglomeration of consumption and production:

$$\lambda - \frac{1}{2} = \frac{2\beta(\alpha - c - \frac{t}{2})}{\gamma N \frac{t}{2}} \left(\theta - \frac{1}{2}\right),\tag{15}$$

which implies that $\lambda > \theta$ as long as there is trade between the two regions $(t < t_{trade})$. This means that the region displaying a higher share of consumption, θ , will attract more than proportional quantity of capital and thus firms, λ . This phenomenon is called the Home Market Effect (HME).

Furthermore, it can be noted that $\frac{\partial \lambda}{\partial t} < 0$, this meaning that lower transport costs, when t is equal in the two directions, always leads to a higher degree of industrial concentration in the regions with the highest level of consumption. A visual representation of this process is provided in Figure 2, where is shown, for every θ , how the range of possible values of λ changes in function of $t \in [\tau; t_{trade}]$, where τ is the cost of supplying the transport service (or alternatively, the lowest possible cost level, which can also be naught and yield corner solutions).

Therefore, the magnitude of the effect is inversely related to the level of transport costs in the two directions, so that a symmetric bilateral decrease in transport costs is always bound to result in a higher degree of industrial concentration in the most affluent region of the economy, and intensify the HME.

In order to give an idea of the outcome interaction between θ and t in the definition of the equilibrium agglomeration of economic activities in the economy, a simulation has been run, where τ has been normalized to unity (which is the prevailing remuneration for labor in the economy) and the other structural parameters have been chosen respecting the restrictions on the domain of certain variable.⁶ The results of the simulation are shown in Figure 3, where each line represents a different level of concentration of consumption in the core region, θ_i . It can be observed, for instance, that even a low level of concentration of consumption, $\theta = 0,55$, associated with very low transport costs can lead to an almost total agglomeration of economic activities in the bigger region and a similar result can be obtained with a much higher concentration, like $\theta = 0,72$, but with the highest transport costs consistent with the existence of inter-regional trade.

To have a further idea of the extent of the Home Market Effect, it suffices to consider that for $t = t_{trade}$, which is the highest value of t compatible with the existence of trade,

⁶For example, β should be bigger than γ for the own substitution effect to be larger than the cross substitution effect.

the following relation holds:

$$\frac{\lambda - \frac{1}{2}}{\theta - \frac{1}{2}} = 2 + \frac{2\beta}{\gamma N}.$$

This means that, even if the highest possible transport cost is considered, defining an imbalance as the difference between the actual concentration of production and consumption, λ and θ , and a perfectly even distribution, $\frac{1}{2}$, it can be seen that imbalance in production are more than twice as big as imbalances in consumption. This can be seen as a lower bound for the Home Market Effect. Higher levels of imbalance depend on structural parameters such as the bias towards the consumption of different varieties, captured by β and γ , and the number of firms in the economy, N, which is a function of the relative importance of the fixed cost of entry.

In order to give an idea of the size of the imbalance magnification effect associated with the changes in transport costs, Table 1 shows the level of production concentration associated with the lowest and the highest levels of transport costs used for plotting Figure 3. It can be remarked that the ratio between imbalance in production and consumption is constant in the level of consumption concentration θ but varies with the transport costs.

θ	$\lambda(t_{trade})$	$\lambda(\tau)$	$ heta-rac{1}{2}$	$\lambda(t_{trade}) - \frac{1}{2}$	$\lambda(au) - rac{1}{2}$
0,5	0,5	0,5	0	0	0
0,51	0,521	0,583	0,01	0,021	0,083
0,55	0,604	0,916	0,05	0,104	0,416
0,6	0,709	1,33	0,1	0,209	0,83
0,65	0,813	1,75	0,15	0,313	1,25
0,7	0,917	2,16	0,2	0,417	1,66
0,75	1,02	2,58	0,25	0,52	2,08
			$\frac{ heta{-1/2}}{\lambda{-1/2}}$:	2,09	8,31

Table 1: Quantification of the HME in the simulation: lower and upper bounds

3.1 The impact of changes in industrial agglomeration on market outcomes

Looking at the price, quantity and profit equations (9), (11) and (5) it can be noted that all the relevant market outcomes are affected by the level of industrial concentration, λ , in the core and in the periphery regions. Before turning to the impact of changes in freight rates on the consumers, it can be useful to analyze the relationship between industrial concentration, λ , and equilibrium prices and quantities, holding the freight rate fixed, as could result for example from an exogenous shock in the relative concentration of consumption, θ .

First of all, from (9) it can be seen that prices in the four segments would react as follows:

$$\frac{\partial p_{HH}}{\partial \lambda} = \frac{\partial p_{FH}}{\partial \lambda} < 0 < \frac{\partial p_{FF}}{\partial \lambda} = \frac{\partial p_{HF}}{\partial \lambda} \quad ; \quad \left| \frac{\partial p_{HH}}{\partial \lambda} \right| = \left| \frac{\partial p_{FH}}{\partial \lambda} \right| = \left| \frac{\partial p_{FF}}{\partial \lambda} \right| = \left| \frac{\partial p_{HF}}{\partial \lambda} \right|$$

The price indices would change accordingly, as can be seen from (8):

$$\frac{\partial P_H}{\partial \lambda} < 0 < \frac{\partial P_F}{\partial \lambda} \quad ; \quad \left| \frac{\partial P_H}{\partial \lambda} \right| = \left| \frac{\partial P_F}{\partial \lambda} \right|.$$

A rising level of concentration of firms in the bigger region is then expected to have opposite and symmetric effects on the consumers' surplus of the two regions. Indeed, in the region hosting more firms tougher competition pushes prices down for both local and imported varieties. The opposite happens in the less industrialized region. Since the total number of firms in the two regions is determined exogenously by the fixed costs of entry, the channel through which competition toughens in the bigger region is the reduction in the number of imported varieties. A lower number of varieties is indeed internalizing the transport costs in their final price, and thus domestic firms need to charge a lower price to remain attractive for their consumers. In the smaller region, exactly the opposite happens: as more goods are imported, the price index rises and firms accurately charge higher prices to maximize their profits.

Similarly the effects of a change in the distribution of economic activities (λ) on quantities sold of each variety can be analyzed from (11):

$$\frac{\partial q_{HH}}{\partial \lambda} = \frac{\partial q_{FH}}{\partial \lambda} < 0 < \frac{\partial q_{FF}}{\partial \lambda} = \frac{\partial q_{HF}}{\partial \lambda} \quad ; \quad \left| \frac{\partial q_{HH}}{\partial \lambda} \right| = \left| \frac{\partial q_{FH}}{\partial \lambda} \right| = \left| \frac{\partial q_{FF}}{\partial \lambda} \right| = \left| \frac{\partial q_{HF}}{\partial \lambda} \right|.$$

Therefore, while changes in the freight rate, t, have been shown in (12) to be distortionary and globally welfare reducing (since as they rise, the rises in quantities of domestic varieties are more the counterbalanced by falls in quantities of imported varieties), any change in the distribution of economic activities is just transferring exactly the same amounts of quantities and price margins from one segment to another, as can be easily seen through this new formulation of (9), (8) and (11). This means that, while reducing transport costs can have positive effects from a global welfare point of view, the extent of spatial distribution is affecting only the relative prosperity of the regions, but not their overall welfare. Thus, as economic imbalance grows between regions, no beneficial effect for the whole economy can compensate for the potentially negative effects of rising inequality in the distribution of economic activities.

It can also be noticed that quantities move in the same direction as prices, which may be surprising at first sight, but it is not once it is realized that the parameter λ affects directly the regional price index.

Evidently, after the introduction on the market of less expensive varieties, whose quality is the same as it was when they were imported but whose price is now free from the burden of transport costs, even the reduction in selling price is not able to help domestic firms keep their production levels as before in the bigger region. Firms exporting in the bigger region then just follow, as their selling volumes are shown in (11) to be systematically lower.

Summing up, from a firm's perspective, after an exogenous increase in the concentration of economic activities, λ , prices and quantities of the increased number of varieties produced in the core region fall in the domestic market and rise in the periphery. The same happens to the fewer firms left in the peripheral region: prices and quantities fall in the core and rise in the domestic market.

It could be interesting to note, at this point, that while the number of varieties and the quantity per variety produced in F and sold in H decrease, the number of varieties and the quantity per variety produced in H and sold in F rise. But what will then be the overall effect on international trade? The answer is not *a priori* clear, as it depends on transport costs and structural parameters in a complex way. The two components of total trade flows (the flow from H to F and the flow from F to H) can be shown to behave asymmetrically and move in opposite directions, i.e. Q_{FH} monotonically decreases in λ whereas Q_{HF} increases. Indeed, the total volume of inter-regional trade in the two directions is determined by three elements: the number of people living in the importing region, M_i , the individual consumption of each imported variety $q_{s,ji}$ and the number of exported varieties (which coincides with the number of firms in the exporting region), N_j . Formally, it is $Q_{ij,ji} = N_{ij}M_jq_{ij} + N_{ji}M_iq_{ji}$, which, plugging (11) and (6), and substituting, can be written as

$$Q_{HF} = \lambda N(1-\theta)Mq_{HF} = \frac{\lambda N(1-\theta)M}{\beta(2\beta+\gamma N)} \left[(\alpha-c)\beta - \frac{t}{2}(2\beta+\gamma(1-\lambda)N) \right]$$

and

$$Q_{FH}(t_{FH}) = (1\lambda)N\theta M q_{FH} = \frac{(1-\lambda)N\theta M}{\beta(2\beta+\gamma N)} \left[(\alpha-c)\beta - \frac{t}{2}(2\beta+\gamma\lambda N) \right].$$

Therefore,

$$\frac{\partial Q_{HF}}{\partial \lambda} \ge 0$$
 and $\frac{\partial Q_{FH}}{\partial \lambda} < 0$ \forall feasible t .

This is so because $\frac{\partial Q_{HF}}{\partial \lambda} \ge 0$ as long as $t < t_{trade} < \frac{\beta(\alpha-c)}{\beta+\gamma N(\frac{1}{2}-\lambda)}$, as can be seen from (14), and thus the condition holds as long as there is trade from region H to region F. The same is true for $\frac{\partial Q_{FH}}{\partial \lambda} < 0$, which holds $\forall t < t_{trade} < \frac{\beta(\alpha-c)}{\beta+\gamma N(\lambda-\frac{1}{2})}$.

This means that an increase in industrial agglomeration increases the shipments from the core to the peripheral region and decreases them from the periphery towards the core. The overall effect can be seen by combining the trade flows in the two directions:

$$Q(t) = Q_{ij,ji}(t) = NM \frac{(\alpha - c)\beta(\theta_i\lambda_j + \theta_j\lambda_i) - t\left[\lambda_i\lambda_j\gamma\frac{N}{2}(\theta_i + \theta_j) + \beta(\theta_i\lambda_j + \theta_j\lambda_i)\right]}{\beta(2\beta + \gamma N)}.$$
 (16)

It is worth noting that the effect is always positive, for every feasible level of interregional transport costs. Formally

$$\frac{\partial Q_{HF,FH}}{\partial \lambda} > 0 \quad \forall t < t_{trade}.$$

4 Introducing the transport sector in the economy

Up to now, the transport cost t in the model has been treated as an exogenous parameter, shaping market outcomes in the differentiated good sector and the location choice of firms. The exogenous definition of t is now replaced by a competitive transport sector in which freight rates are determined endogenously.

In other words, the transport sector is introduced in the economy, turning the exogenously given transport costs into endogenously determined freight rates. This analysis then feeds into the next section, where it is analyzed how changes in the freight rate affect market outcomes and shape the economic geography of the regions. Two different configurations of the transport sector are considered, each one corresponding to a different type of liberalization:

- First, the case is considered of a transport sector in which carriers are allowed only to charge the same price per route or, equivalently, to set the tariff based only on distance and not on the direction of the shipment. This is referred to as the **unique-bilateral-tariff liberalization**;
- Second, the case is considered of a transport in which carriers are left free to set their prices freely and possibly segment their market based on the direction of the shipment (which is the only source of heterogeneity among their clients). This is referred to as **complete deregulation**.

The central difference between the two regimes is, thus, the possibility of segmenting the transport market into HF and FH submarkets.

4.1 Market structure in the transport sector

Before showing how the transport sector is actually modeled, it will be useful to spend a few words on the nature of the transport service to justify the modeling strategy. First of all, it should be decided whether the transport service should be considered within the perfect or the imperfect competition framework. As a matter of fact, it can be claimed that there are several dimensions over which firms operating in the transport sector (from now on, carriers) can differentiate their service, both vertically (speed, punctuality, traceability and so on) or horizontally (specializing in particular sector of the economy or geographical areas). Yet, none of this would be conceivable in an economy characterized by identical consumers and identical firms producing a continuum of horizontally differentiated varieties of the same good in four segments, whose only distinctive features are the regions of production and sale.

This consideration, pushes in the direction of modeling the transport sector in a more classical way, treating the transport service as homogeneous. This said, if also free entry was assumed, prices would be pushed to marginal costs and the contribution of the transport sector to the dynamics of the model would be null. The most straightforward way to model the transport sector is then to turn to a standard oligopoly, in which a fixed number of firms, K, compete in quantities or, equivalently, compete in prices after having committed to a certain capacity (Vives, 1999). This would be a reasonable assumption, as transport's capacity and infrastructures have to be built before the service is effectively sold. This market structure is also convenient in terms of analytical tractability, as the only decision available to regulators to shape the market outcomes is to decide the numbers of competitors to allow in the transport sector.⁷

4.2 Endogenization of the transport costs under a unique tariff

Assuming that all the carriers have the same cost structure, the profit function determining the k^{th} carrier's behavior resulting from the market structure adopted for the transport sector can be written as

$$\Pi_k = (t - \tau)q_k,\tag{17}$$

where t is the freight rate, τ is the marginal cost of transport and q_k represents the quantity of goods delivered by each carrier, $q_{k,ij} = Q_{ij}/k$. Notice that the expression capturing the total amount of international trade, as computed in equation (16), can now be seen as the demand function associated with the transport sector. It is well behaved in the sense that it linearly decreases in the parameter of interest, the freight rate t, so that the carriers' problem is well defined and yields an interior solution: as the price of the transport service rises, the imported goods become more expensive and the inter-regional trade flows decline. More formally, it can be shown that the elasticity of international

⁷For simplicity, without loss of generality, the number of firms in the transport sector is chosen by the regulator. In alternative, the amount of fixed entry costs could be decided by the regulator and indirectly determine the number of carriers associated with a sufficient level of profits to cover the fixed costs.

trade to transport costs is negative and decreasing in t.⁸

$$\epsilon_{Qt} = -\frac{\partial Q}{Q} \cdot \frac{t}{\partial t} \equiv -\frac{\partial Q}{\partial t} \cdot \frac{t}{Q}$$

$$\Rightarrow \ \epsilon_{Qt} = \frac{t[\theta_i \lambda_j (2\beta + \gamma N \lambda_i) + \theta_j \lambda_i (2\beta + \gamma N \lambda_j)]}{2(\alpha - c)\beta(\theta_i \lambda_j + \theta_j \lambda_i) - t[\theta_i \lambda_j (2\beta + \gamma N \lambda_i) + \theta_j \lambda_i (2\beta + \gamma N \lambda_j)]},$$

where $\epsilon_{Qt} \in [0; +\infty] \ \forall t \in [0; t_{trade}]$. A visual representation of the carrier's problem is provided in Figure 4, where it can be observed that the problem of the carriers closely resembles the problem of suppliers of an intermediate input in the context of a double marginalization setting (Tirole, 1988).

In order to solve analytically the carrier's problem, it is convenient to express the freight rate, t, in terms of the inter-regional flow of goods, given by (16), and then maximize profit function (17) with respect to total quantities shipped. The inverse demand for inter-regional transport is then

$$t(Q) = \frac{(\alpha - c)\beta(\theta_i\lambda_j + \theta_j\lambda_i) - Q\frac{\beta(2\beta + \gamma N)}{NM}}{\lambda_i\lambda_j\gamma\frac{N}{2}(\theta_i + \theta_j) + \beta(\theta_i\lambda_j + \theta_j\lambda_i)}$$

and the resulting optimization problem of each carrier results in the pricing function

$$t^* = \frac{k}{k+1}\tau + \frac{2(\alpha - c)\beta(\theta_i\lambda_j + \theta_j\lambda_i)}{(k+1)[\theta_i\lambda_j(2\beta + \gamma N\lambda_i) + \theta_j\lambda_i(2\beta + \gamma N\lambda_j)]}.$$
(18)

This expression has the advantage of being fairly general. In fact, ranging from K = 1 to $K \to \infty$, all the possible market structures involving homogeneous goods are met. This latter case, indeed, corresponds to a perfect competition framework where $t^* \to \tau$, the marginal cost of production and profits are null. The opposite is true for the former case, where K = 1 and the monopolist can take advantage of its market power and extract the maximum possible level of profits, charging the typical price a monopolist charges when

$$\frac{t^* - \tau}{t} = \frac{1}{k} \cdot \frac{1}{\epsilon_{Qt}}$$

⁸The analysis of the elasticity is rather convenient too in the study of monopolies and oligopolies. Indeed, the multiplicative inverse of the price elasticity equals exactly the relative mark-up which a pure monopolist will charge in the market. This result is easily extendable to oligopolists by simply dividing this value by the number of competitors. For example, in the case here analyzed, it would hold the following relationship:

And so, roughly speaking, doubling the elasticity would have exactly the same effects on oligopolistic prices as doubling the number of firms.

facing a linear demand function: half of the maximum price compatible with positive amount of service sold, plus half of any marginal cost of production on which it incurs.

As the system gets further away from a monopoly, prices fall: in other words, t^* declines as the number of competitors, K, increases. Regulators can thus indirectly determine the resulting price in the transport sector by just setting the number of competitors.

4.3 Endogenization of the transport costs under segmentation

In a completely deregulated regime, profit maximizing carriers are expected to segment their markets because all the conditions for segmenting are satisfied: their revenues will be higher, consumer screening is costless and arbitrage between consumers (firms involved in inter-regional trade, in this case) is not profitable.

The first condition can be inferred by the analysis of the elasticities of trade flow to the freight rate in the two segments. The elasticity in the segment from F to H will be shown to be higher than the one from H to F, in equations (28) and (27). This means that applying the same price to both segments means to charge a too high price in the first or a too low in the second (or both, as will actually happen when a unique freight tariff has to be chosen) and profits will not be maximized.

The second condition holds because the key characteristic for segmentation is impossible to hide, since the very purchase of the transport service in one region rather than the other gives information about the segment which is being served.⁹

Finally arbitrage can be excluded, as the third condition states, because it is never profitable to carry it out as can be argued from the following relationship, based on equations (9) defining equilibrium prices:

$$p_{FF}^* - p_{HH}^* = \frac{t}{2} \cdot \frac{\gamma N(\lambda_H - \lambda_F)}{(2b + \gamma N)} < t.$$

This means that no third agent could make profits out of buying a good in one region and reselling it in the other one, as long as some transport service has to be purchased.

⁹It can also be noted that the assumption of identical firms (or, at least, technologies of production) rules out the possibility of alternative, cheaper ways of getting from one region to the other: if even one firm were able to deliver its products to the other region in a cheaper way, indeed, all the other firms would be also able to, leaving no room for the existence of a specific transport sector.

4.3.1 Redefining economic variables under market segmentation

Once segmentation in the transport market is allowed, the relationships between economic variables have to be rewritten to take into account of the potentially different transport costs in the two directions of shipment. Here follows the description of the economy under a segmented transport sector.

Consumer demand

The formal representation of consumers' utility function is essentially unaffected and so is the resulting demand function. The only difference with respect to the demand function (3) is that the prices entering the demand function are now different, as they depend on the segmented freight rates t_{ji} and no more on a unique t, as previously assumed:

$$q_{s,ji}(t_{ji}) = \frac{\alpha - \gamma Q_i(t_{ji})}{\beta} - \frac{p_{s,ji}(t_{ji})}{\beta} = \frac{\alpha\beta + \gamma N\bar{p}_i(t_{ji})}{\beta(\beta + \gamma N)} - \frac{p_{s,ji}(t_{ji})}{\beta}$$

Firms

In order to reflect the fact that exporters of the two regions are now facing different transport costs, profit functions have to be slightly adjusted:

$$\Pi_{s,H} = (p_{s,HH} - c)q_{s,HH}M_H + (p_{s,HF} - c - t_{HF})q_{s,HF}M_F - r_Hf$$

and

$$\Pi_{s,F} = (p_{s,FF} - c)q_{s,FF}M_F + (p_{s,FH} - c - t_{FH})q_{s,FH}M_H - r_F f$$

Also profit-maximizing prices have to be adapted:

$$p_{s,HH}(t_{FH}) = \frac{\alpha\beta + c(\beta + \gamma N) + \gamma N\lambda_F \frac{t_{FH}}{2}}{2\beta + \gamma N} \quad ; \quad p_{s,FH}(t_{FH}) = p_{s,HH} + \frac{t_{FH}}{2} \tag{19}$$

and

$$p_{s,FF}(t_{HF}) = \frac{\alpha\beta + c(\beta + \gamma N) + \gamma N\lambda_H \frac{t_{HF}}{2}}{2\beta + \gamma N} \quad ; \quad p_{s,HF}(t_{HF}) = p_{s,FF} + \frac{t_{HF}}{2}. \tag{20}$$

Thus, the price index can be recomputed as follows:

$$\bar{p}_H(t_{FH}) = \frac{P_H(t_{FH})}{N} = \frac{\alpha\beta + (c + \lambda_F t_{FH})(\beta + \gamma N)}{2\beta + \gamma N}$$
(21)

and

$$\bar{p}_F(t_{HF}) = \frac{P_F(t_{HF})}{N} = \frac{\alpha\beta + (c + \lambda_H t_{HF})(\beta + \gamma N)}{2\beta + \gamma N}.$$
(22)

An important aspect illustrated by this formulation of the price indices is that, as long as t is fixed in the two directions, the price index in H is always lower than in F, but this may not necessarily be the case when carriers segment the transport market. Finally, given these prices and price indices, the equilibrium quantities sold in the two markets become:

$$q_{s,HH}(t_{FH}) = \frac{(\alpha - c)\beta + \gamma N\lambda_F \frac{t_{FH}}{2}}{\beta(2\beta + \gamma N)} \quad ; \quad q_{s,FH}(t_{FH}) = \frac{(\alpha - c)\beta - \frac{t_{FH}}{2}(2\beta + \gamma N\lambda_H)}{\beta(2\beta + \gamma N)} \tag{23}$$

and

$$q_{s,FF}(t_{HF}) = \frac{(\alpha - c)\beta + \gamma N\lambda_H \frac{t_{HF}}{2}}{\beta(2\beta + \gamma N)} \quad ; \quad q_{s,HF}(t_{HF}) = \frac{(\alpha - c)\beta - \frac{t_{HF}}{2}(2\beta + \gamma N\lambda_F)}{\beta(2\beta + \gamma N)}.$$
(24)

4.3.2 The segmented transport sector

The study of the transport sector's dynamics will proceed as before. Carriers' profits maximizing behavior is reflected in the same way as in (17), with the only difference that now the inter-regional flows in the two direction are considered and used as the demand function for the segmented transport services.

The two components of aggregate volume of inter-regional trade flows will then be $Q_{HF}(t_{HF})$ and $Q_{FH}(t_{FH})$, whose sum then represents the whole flow. Formally it is

$$Q = Q_{HF}(t_{HF}) + Q_{FH}(t_{FH}),$$

where

$$Q_{HF}(t_{HF}) = \lambda N(1-\theta)Mq_{HF} = \frac{\lambda N(1-\theta)M}{\beta(2\beta+\gamma N)} \left[(\alpha-c)\beta - \frac{t_{HF}}{2}(2\beta+\gamma(1-\lambda)N) \right]$$
(25)

and

$$Q_{FH}(t_{FH}) = (1-\lambda)N\theta M q_{FH} = \frac{(1-\lambda)N\theta M}{\beta(2\beta+\gamma N)} \left[(\alpha-c)\beta - \frac{t_{FH}}{2}(2\beta+\gamma\lambda N) \right].$$
(26)

This means that the transport services offered in the two segments are now traded in different markets, each one characterized by a specific elasticity of inter-regional trade flows to transport cost:

$$\epsilon_{Q_{HF}t_{HF}} = t_{HF} \frac{(2\beta + \gamma(1-\lambda)N)}{2(\alpha - c)\beta - t_{HF}(2\beta + \gamma(1-\lambda)N)}$$
(27)

and

$$\epsilon_{Q_{FH}t_{FH}} = t_{FH} \frac{(2\beta + \gamma\lambda N)}{2(\alpha - c)\beta - t_{FH}(2\beta + \gamma\lambda N)},\tag{28}$$

which are both increasing in the freight rate, as expected, but evolve in opposite directions as λ changes. Indeed it can be verified that

$$\frac{\partial \epsilon_{Q_{HF}t_{HF}}}{\partial \lambda} > 0 \quad ; \quad \frac{\partial \epsilon_{Q_{FH}t_{FH}}}{\partial \lambda} < 0.$$

From (25) and (26) it is then possible to derive the two inverse demand functions, which are

$$t_{HF} = \frac{1}{2\beta + \gamma(1-\lambda)N} \left[2(\alpha - c)\beta - \frac{\beta(2\beta + \gamma N)}{\lambda N(1-\theta)M} Q_{HF} \right]$$

and

$$t_{FH} = \frac{1}{(2\beta + \gamma\lambda N)} \left[2(\alpha - c)\beta - \frac{\beta(2\beta + \gamma N)}{(1 - \lambda)N\theta M} Q_{HF} \right].$$

Then, plugging them into equation (17) and optimizing with respect to quantities, the prevailing freight rates on the two segments become

$$t_{HF} = \frac{k}{k+1}\tau + \frac{2(\alpha - c)\beta}{(k+1)(2\beta + \gamma(1-\lambda)N)} \quad \Rightarrow \quad t_{HF} > t$$
⁽²⁹⁾

and

$$t_{FH} = \frac{k}{k+1}\tau + \frac{2(\alpha - c)\beta}{(k+1)(2\beta + \gamma\lambda N)} \quad \Rightarrow \quad t_{FH} < t.$$
(30)

Comparing these two freight rates in the segmented markets with the equation in (18), which describes the freight rate carriers would choose for the transport service if not allowed to price discriminate, it can be noticed that transport from the smaller to

the bigger region will be cheaper than under the unique tariff, but transport from the bigger to the smaller would be more expensive:

$$t_{FH} < t < t_{HF}.\tag{31}$$

Indeed, the unique-bilateral-tariff pricing equation shown in (18) turns out to be a weighted average of the two segmented tariffs (29) and (30). This means that, moving from a regulated tariff system, implying a unique freight rate in both directions, to a deregulated one, leading to segmentation, is likely to affect in a complex way all the agents of the economy: inter-regional delivery would be cheaper from F to H, but more expensive from H to F and this affects local consumers and firms in opposite ways, limiting the agglomeration in the core region.

In the next section a welfare analysis is undertaken on the impact of transport regime's change in the two regions. As will be shown, gainers and losers can be identified in each region, but to different extents.

5 Comparing regulatory regimes

In this section, the two kinds of liberalization are considered. The starting point is an economy with a very uncompetitive transport market, before regulatory intervention, characterized by monopolistic or high oligopolistic prices. Two policy options are then available to public authorities willing to reduce freight rates by liberalizing the transport sector:

- A regulated **unique-bilateral-tariff liberalization**, imposing carriers to charge the same price per route or, equivalently, to set the tariff based only on distance and not on the direction of the shipment;
- A complete deregulation, allowing carriers to set their prices freely and possibly segment their market based on the direction of the shipment (which is the only source of heterogeneity among their clients).

Trade liberalization regimes are compared in a sequential way. First, the transition from an expensive and heavily regulated framework to a unique-bilateral-tariff liberalization is considered, as captured by an increase in the number of carriers. Then, the additional effect of allowing carriers to price discriminate by the direction of shipment is explored.

5.1 Liberalization under unique tariff

The focus of the welfare analysis is on the effects of transport liberalization on prices, quantities and industrial agglomeration in the heterogeneous good sector.

Prices

As for prices, equation (9) implies that that prices of the heterogeneous goods decrease together with transport costs, but with different intensities in the different segments. Indeed, assuming a unique tariff and looking at the first derivatives of (9) (or, equivalently, (19) and (20), holding $t_{ij} = t_{ji}$) with respect to t, it suffices to remind that $\lambda \equiv \lambda_H > \lambda_F$ to see that

$$\frac{\partial p_{HF}}{\partial t} > \frac{\partial p_{FH}}{\partial t} > \frac{\partial p_{FF}}{\partial t} > \frac{\partial p_{HH}}{\partial t} > 0.$$

A reduction in transport costs thus affects all the manufacturing firms in the economy, but while in the domestic segment the sales of the F-located firms are more affected than H-located firms, in the export segments the situation is reversed. From a consumers' standpoint this means also that prices in F will decrease more than prices in H: this is a consequence of the higher share of imported varieties in the smaller market. Indeed, it can also be checked from (21) and (22) that

$$\frac{\partial \bar{p}_F}{\partial t} > \frac{\partial \bar{p}_H}{\partial t},$$

which implies higher consumer surplus gains in the periphery than in the core.

Quantities

The resulting quantities dynamics have been shown in (12) to be

$$\frac{\partial q_{HH}}{\partial t} > 0 \quad ; \quad \frac{\partial q_{FH}}{\partial t} < 0$$

and

$$\frac{\partial q_{FF}}{\partial t} > 0 \quad ; \quad \frac{\partial q_{HF}}{\partial t} < 0.$$

Note that, interestingly, movements in opposite directions have been found also as a consequence of changes λ . However there is an important difference between the effects of variations in the parameters t and λ : the latter, in fact, implies a zero-sum transfer of quantities sold, i.e. a perfectly balanced and symmetric variation. However, this is not the case for transport costs, whose variation yields

$$\left|\frac{\partial q_{ii}}{\partial t}\right| < \left|\frac{\partial q_{ji}}{\partial t}\right|. \tag{32}$$

This relationship shows why transport costs can be considered intrinsically distortionary and why their reduction is unambiguously found to be welfare improving in the short run.

Agglomeration and Manufacturing firms' Operating Profits

As for the agglomeration of economic activities, it can be see from equation (15) a reduction in transport costs implies a magnification of the disparities in the long run:

$$\frac{\partial \lambda}{\partial t} < 0.$$

This reallocation of resources is triggered by a disproportionate impact of changes in t on the firms based in the smaller region. It may appear counterintuitive, as firms in the peripheral region have to incur lower transport costs to serve customers in the other region. However, this effect has to be traded off against tougher competition coming from the firms located in the core region, which now have easier access to the peripheral markets. Equation (15) is derived from the equalization of capital remuneration across the two regions and it signals that profits of firms located in the smaller region are affected more severely than those in the bigger region from the intensification of competition due to lower transport costs, so as to lead a higher share of region F's capital to flow toward H and a higher relative number of varieties produced there.

In the short run (i.e. before capital remuneration re-equalize between regions), firms' profits in the different segments are affected in a complex way. Indeed, two interacting effects are at play: profits in the export segments rise as a consequence of higher mark-ups (remember that firms pass through half of their linear transport costs) and quantities sold. Yet, operating profits on the domestic market fall because tougher competition lowers

prices and quantities sold of each variety, i.e.:

$$\frac{\partial \Pi_{HH}}{\partial t} > 0 \quad ; \quad \frac{\partial \Pi_{FF}}{\partial t} > 0$$

and

$$\frac{\partial \Pi_{FH}}{\partial t} < 0 \quad ; \quad \frac{\partial \Pi_{HF}}{\partial t} < 0.$$

The overall effect again depends on structural parameters (such as the relative market sizes, θ_i , and the bias toward a varied consumption, c) but the HME relationship (15) ensures us that firms in the bigger region systematically outperform firms in the smaller one in the short run, before capital returns are equalized again across the regions:

$$\frac{\partial \Pi_H}{\partial t} > \frac{\partial \Pi_F}{\partial t} > 0.$$

Thus they lose less or earn higher profits than smaller region's ones as a consequence of liberalization.

Carriers' profits

If we consider the previously regulated transport sector as behaving like a monopoly or, at least, we assume it to be a less competitive (because of fewer players) oligopoly than after the liberalization takes place, the global level of profits generated by the transport sector decreases because of a higher intensity of competition between carriers.

However, the existence of some residual market power (as long as $k \to \infty$ or $t \to \tau$) still causes some welfare dead-weight losses in the economy because every inter-regionally traded product will still be incorporating both the mark-up applied by carriers for their transport service and the mark-up applied by manufacturing firms. The former is caused by the oligopolistic structure of the transport sector and the latter is due to the market power conferred by the horizontal differentiation. As observed earlier, this phenomenon mirrors what is usually called *double marginalization* in industrial organization and is known to lead to an under-provision of good from both a consumers' and a manufacturing firms' standpoints (Tirole, 1988).

Global welfare

The analysis above shows how, at least in the short run, liberalization (such as any other shock or policy whose result is to reduce the transport costs) is welfare improving for the entire economy. The result is mainly driven by the evolution of prices, which fall in all the four segments of the heterogeneous goods' sector, as the transport costs fall.

An important remark is that the distribution of economic activities gets more unbalanced while global welfare rises, which creates a clear trade-off for policy makers. Nonetheless, welfare improvements in the deindustrializing region, F, are stronger than in the bigger region, because price reductions are more intense. However, from (8) (or, equivalently, from (21) and (22) when $t_{HF} = t_{FH}$) it can be seen that, when t is identical in the two directions of trade, F dwellers could never catch up entirely with the welfare level of consumers living in H, in the absence of proper redistributive mechanisms.

5.2 Complete deregulation: the economic impact of transport segmentation

The analysis of full deregulation leads to rather different conclusions. The departure from the symmetric transport cost towards a direction-specific pricing implies that uniquebilateral-tariff liberalization and full deregulation have different impacts on regional prices. In particular, as it has been noted, the symmetric outcome shown in equation (18) is the weighted average of equations (29) and (30), where weights are represented by the relative number of firms and consumers in the two regions. Therefore moving from a unique-bilateral-tariff liberalization to full deregulation, transport prices in the two directions move in opposite directions.

Prices

The differentials on prices can be easily deducted by the analysis of (19) and (20), keeping in mind the relationship stated in (31). They can be summarized as follows:

$$p_{HH}^S < p_{HH} \quad ; \quad p_{FH}^S < p_{FH}$$

and

$$p_{FF}^S > p_{FF} \quad ; \quad p_{HF}^S > p_{HF}$$

where the superscript S stands for market outcomes after carriers are allowed to segment the transport market. What is worth noting is that prices increase for consumers in the smaller region for both domestic and imported goods. The opposite holds for consumers in the bigger region. This results in price indices diverging further, as can be checked using equations (21) and (22):

$$\bar{p}_{H}^{S}(t_{FH}) < \bar{p}_{H}(t) \; ; \; \bar{p}_{F}^{S}(t_{HF}) > \bar{p}_{F}(t).$$

This effect reinforces the relation noted in (8) and yields

$$\bar{p}_{H}^{S}(t_{FH}) < \bar{p}_{H}(t) < \bar{p}_{F}(t) < \bar{p}_{F}^{S}(t_{HF})$$

Segmentation is then expected to increase the gap in prices between the two regions as compared to a unique tariff in the two directions.

Quantities

More complex is the impact of segmentation on quantities, as can be derived from equations (23) and (24), taking into account the effect on prices shown in (31):

$$q_{HH}^S < q_{HH} \quad ; \quad q_{FH}^S > q_{FH}$$

and

$$q_{FF}^S > q_{FF} \quad ; \quad q_{HF}^S < q_{HF}$$

These results show how, after segmentation, each variety produced by the firms located in region F would sell more than before in terms of quantities, both in the local and the export segments. The opposite will happen to the firms located in H. Hence, compared to a unique-bilateral tariff, complete deregulation implies higher levels of production in Fand higher employment in the manufacturing sector. In addition, given the higher level of domestic prices, complete deregulation engender higher profits in the short run to the firms located in F, as compared to a unique-bilateral tariff. This latter effect plays against agglomeration and yields the following result.

Agglomeration and Manufacturing Firms' Operating Profits

Moving from a unique-bilateral-tariff liberalization to complete transport deregulation, i.e. once carriers are allowed to price discriminate by direction of shipment in the transport sector, no clear trend can be seen on agglomeration anymore. Equation (15) has indeed to be generalized into

$$\lambda - \theta = \frac{2\beta(\alpha - c)}{\gamma N} \left(\frac{\theta}{t_{FH}} - \frac{1 - \theta}{t_{HF}} \right) - \left(\frac{\beta + \frac{\gamma N}{2}}{\gamma N} \right) \left(\frac{\theta t_{HF}}{t_{FH}} - \frac{(1 - \theta)t_{FH}}{t_{HF}} \right), \quad (33)$$

where capital remuneration in the two regions is now expressed in terms of two different freight rates and the changes in profitability of firms in the two regions depend on structural parameters and levels of concentration of consumption, θ , in a non-linear way. From a spatial agglomeration standpoint, then, the interaction between mobile returnsmaximizing capital and price-discriminating carriers does not yield so clear-cut results as in New Economic Geography settings in which t is treated parametrically or is not allowed to vary by the direction of shipment. In fact, *industrial agglomeration may even turn out to be lower than consumption agglomeration if the number of firms (N) or the substitutability between varieties (\gamma) are high enough.*

This result can be verified by looking at the right-hand side of (33) and noticing that for $\gamma N \to \infty$ the first term tends to zero. However, the opposite is true when the preference for the differentiated good increases (α) or the marginal cost of production (c) decreases. As a matter of fact, the first term on the right-hand side of (33) is always positive because $\theta > 1 - \theta$ by definition and $t_{FH} < t_{HF}$ as shown in (31). For the same reasons, the second term on the right-hand side of (33) is always negative. The relative strength of one term over the other is then basically determined by the intensity of the economies of scale: higher α and lower β , N and γ all contribute to increasing the markups of vis-à-vis the fixed costs of entry and favor agglomeration, i.e. a higher λ .¹⁰

The short-run analysis of how profits are affected by segmentation in the transport sector is complicated by the fact that, even if quantities rise for every segment supplied by firms located in F and falls fall for varieties produced in H, selling prices of each manufacturing firm move in different directions in the domestic and the export segments,

¹⁰It is also interesting to notice that, as expected, when $\theta = 1/2$ and thus $t_{HF} = t_{FH}$, the right-hand side of (33) disappears and with it the Home Market Effect, i.e. $\lambda = \theta$.

which are populated a different share of the total population. Therefore, even if it surely holds that

$$\Pi_{FF}^S > \Pi_{FF} \quad ; \quad \Pi_{HH}^S < \Pi_{HH},$$

there no clear trend for Π_{HF} and Π_{FH} , because in the FH segment sales are higher but prices lower and in the HF case the opposite holds. Thus, since the magnitude of these effects depend on the different parameters of the model, these results help explain the undefined pattern underlined in the changes in agglomeration.

5.2.1 Global welfare

In terms of welfare, accounting for the possibility of carriers to price discriminate across segments has significant implication on the distribution of gains from transport liberalization and on the impact on agglomeration.

As for the welfare gains, segmentation increases consumer gains in the core region by decreasing prices even further than a unique-bilateral tariff. Yet, this downward pressure on prices in the larger region squeezes profits in the domestic segment of the manufacturing firms located in H. This effect drives capitals away from H and poses some restraint on the Home Market Effect pattern triggered by the reduction of transport costs, thus retaining more production in F.

However, since workers are assumed to be always able to participate in the homogeneous good's sector and this rules out any concerns about unemployment in the model, the potential concern is that segmentation transfers to H part of the welfare of consumers in F, without necessarily improving global welfare. On the contrary, it is likely to reduce it as a consequence of higher profits in the transport sector due to the possibility of price discriminating their clients. This means that, whereas moving from a regulated to a liberalized regime is expected to be welfare improving for everyone, moving from a unique-bilateral-tariff regime to complete deregulation may raise redistributive concerns.

Complete deregulation thus turns out to be more beneficial for the larger region, which is already expected to receive capital inflows due to the Home Market Effect associated with lower transport costs. However, it should be noted that any kind of transport liberalization reduces prices in both regions and, notably, more in the periphery than in the core, as can be seen from (32). This means that the smaller region has more to gain in the event of liberalization than the bigger region, notwithstanding the capital outflow. Therefore, depending on whether the preferences of policy makers are more biased toward equalizing consumer welfare or the geographic distribution of economic activities, either full deregulation or a unique-bilateral-tariff liberalization can be pursued.

These results and trade-offs are summarized in Table 2, showing how, from a welfare point of view, both liberalization regimes are to be preferred to a regulated or uncompetitive transport sector. This result can be easily generalized to any policy aimed at reducing freight rates. In fact, prices will fall in the two regions, as can be seen from (21) and (22), and the total global production and employment will rise, as can be seen from (32) and (32), taking (32) into account and bearing in mind that the total number of firms, or equivalently varieties, in the economy is held fixed. Besides the interregional redistributive issue, the only additional source of concern, in terms of policy, may be that the imbalance in the spatial distribution of economic activities reach any socially undesirable level.

	Regulated transport	Liberalized transport sector	
	Fixed freight rate	Unique Tariff	Full Deregulation
H Consumers' surplus	III best	II best	I best
F Consumers' surplus	III best	I best	II best
Industrialization in H	III best	II best	I best
Industrialization in F	III best	I best	II best
Carriers' profits	I best	III best	II best
Restraint of HME	I best	III best	II best

Table 2: Summary of gainers and losers from different transport regimes

Hence, in terms of policy implications, different choices could be made depending on the policymaker's priorities. If the main objective is to reduce the gap in welfare levels cross regions, probably a unique-bilateral-tariff liberalization is to be preferred, as, under complete deregulation, equations (21) and (22) warrant that price differences in the two regions would be larger. Differently, if the aim of political action is to keep the employment levels in the heterogeneous goods' manufacturing sector as evenly distributed as possible, then a complete deregulation is to be opted for.

6 Conclusions

The present analysis confirms that liberalization in the transport sector is expected to yield static welfare gains for consumers of the entire economy. In the short run, tougher competition in the manufacturing sector is induced in each regional market as a result of the cheaper inter-regional connections. This pushes prices down in every segment of the manufacturing good's sector and leads to higher inter-regionally traded quantities of each variety and lower prices

However, the geographic allocation of consumers' gains between the core and the peripheral region depends heavily on the liberalization regime chosen. In fact, when carriers are left free to segment their markets consumers of the larger region would benefit from an even higher share of these gains, if compared with the circumstance in which a unique tariff is imposed in the two directions of trade. Another implication is that the number of people working in the manufacturing sector is expected to decrease less in the smaller region. This means that different policies can be adopted depending on the actual priorities of the political agenda: a unique-bilateral-tariff liberalization should be preferred if consumers' welfare equalization has the highest priority; differently, a complete deregulation should be selected if production and employment in the manufacturing sector have to be kept as evenly distribute as possible, but without renouncing to the welfare gains associated with liberalization.

The empirical relevance of the theoretical results of the model in the context of complete deregulation are warranted by the evidence that prices do differ for shipments in different directions of the same route in the absence of ad hoc regulation. For example Tanaka and Tsubota (2014) observe significant directional price differences using microlevel data on prices charged by carriers connecting different Japanese prefectures. The same observation is made by Kleinert and Spies (2011).

Finally, at least three remarks are due. First of all, in the present model only interactions between manufacturing and transport sectors have been analyzed, thus nothing can be easily inferred about liberalization processes affecting only the transport of commuters or travelers. Second, the implicit simplifying assumption on which this work has relied is that the most direct effect of liberalizations is to reduce prices. Evidently this is not always true, but this doesn't invalidate the underlying analysis, which can be extended to any shock affecting transport costs. It is indeed worth noting that the results here obtained are much more general than that and can also be applied to other variables such as efficiency gains in transport derived from technological improvement or, conversely, inefficiencies generated by higher marginal costs of delivery (for example, trade tariffs, fuel price and so on). An interesting analysis of the evolution of transport costs, underlying these aspects and others, has been conducted by Hummels (2007), who finds out that ad-valorem prices of ocean shipping and air shipping displayed sensibly different trends in the last fifty years. Whereas the latter declined utterly because of technological progress, the former kept constant as a consequence of increasing prices in upstream markets.

In this paper a purely theoretical model has been presented. A natural next step, as an avenue for future avenues of research, is to test empirically some of the implications of the model, especially concerning the determination of the freight rate in the transport sector.

References

- Behrens K., Gaigné C. and Thisse J.-F. (2009), Industry location and welfare when transport costs are endogenous. *Journal of Urban Economics* 65(2), 195-208.
- Behrens, K. and Picard, P. (2011), Transportation, freight rates, and economic geography. Journal of International Economics 85(2),-184 280-291.
- Belleflamme P., Picard P. and Thisse J.-F. (2000), An Economic Theory of Regional Clusters. Journal of Urban Economics 48(1), 158-184.
- Brander, J. and Spencer, B. (1984), Trade warfare: Tariffs and cartels. Journal of International Economics 16(3-4), 227-242.
- Cournot A. A. (1838), Researches into the Mathematical Principles of the Theory of Wealth; English Translation, MacMillan, London, England, 1897.
- Dixit, A. and Stiglitz, J. (1977), Monopolistic Competition and Optimum Product Diversity. American Economic Review 67, 297-308.
- European Commission (2001), White Paper European transport policy for 2010: time to decide, Office for Official Publications of the European Communities, Luxembourg.

- Fujita M. (1988), A monopolistic competition model of spatial agglomeration: A differentiated product approach. *Regional Science and Urban Economics* 18, 87-124.
- Hummels D. (2007), Transportation Costs and International Trade in the Second Era of Globalization, Journal of Economic Perspectives 21, 131-154.
- Kleinert, J. and Spies, J. (2011), Endogenous Transport Costs in International Trade. IAW Discussion Papers 74.
- Ottaviano G.I.P., T. Tabuchi and J.-F. Thisse (2002), Agglomeration and trade revisited. International Economic Review 43, 409-436.
- Picard P. and Zeng D.-Z. (2005), Agricultural sector and industrial agglomeration. Journal of Development Economics 77(1), 75-106.
- Tanaka K. and Tsubota K. (2014), Agglomeration and directional imbalance of transport costs: the role of density economies, *forthcoming*.
- Tirole J. (1988), The Theory of Industrial Organization, MIT Press, Cambridge MA.
- Vives X. (1999), Oligopoly pricing: old ideas and new tools, MIT Press, Cambridge, MA.



Figure 1: Strategic complementarity between $p_{s,ii}$ and \bar{p}_i .



Figure 2: The Home Market Effect. On the left pane is represented the intuition behind the agglomeration measure. On the right pane is shown the range of possible values of λ , depending on the values of t and θ .



Figure 3: Simulation of changes in the agglomeration of industrial activities, λ , as a function of transport costs, t, and agglomeration of consumption, θ .



Figure 4: Visual illustration of how final good's demand structure affects the carriers' problem.