ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Ottaviano, Gianmarco I. P.

Working Paper

Footloose capital, market access, and the geography of regional state aid

HWWA Discussion Paper, No. 132

Provided in Cooperation with:
Hamburgisches Welt-Wirtschafts-Archiv (HWWA)

Suggested Citation: Ottaviano, Gianmarco I. P. (2001) : Footloose capital, market access, and the geography of regional state aid, HWWA Discussion Paper, No. 132, Hamburg Institute of International Economics (HWWA), Hamburg

This Version is available at:
https://hdl.handle.net/10419/19419

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

HWWA HAMBURG

DISCUSSION PAPER

Footloose Capital, Market Access, and the Geography of Regional State Aid

Gianmarco I.P. Ottaviano

HWWA DISCUSSION PAPER

132

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
Hamburg Institute of International Economics
2001
ISSN 1616-4814

The HWWA is a member of:

- Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)
- Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute (ARGE)
- Association d'Instituts Européens de Conjoncture Economique (AIECE)

Footloose Capital, Market Access, and the Geography of Regional State Aid*

Gianmarco I.P. Ottaviano[†]

[†]Università Commerciale'L. Bocconi' Milan, GIIS Geneva and CEPR London

*Preliminary draft (6 May 2001). I am indebted with seminar at University of Bologna Centre in Buenos Aires, University of Milan Bicocca, Catholic University Milan, HWWA Hamburg, NHH Bergen as well as CEPR/RTN Workshop on 'The Economic Geography of Europe' in Villars for comments on earlier drafts. In particular, I wish to thank Fabrizio Barca, André Sapir and Jacques Thisse for discussions on policy implications as well as Gabriele Tondl and Thierry Mayer for the provision of relevant data. Financial support from Bocconi University, University of Bologna, HWWA and the European Commission is gratefully acknowledged.

The subject of this paper is assigned to the HWWA's research programme "European Integration and Spatial Development".

HWWA DISCUSSION PAPER

Edited by the Department
EUROPEAN INTEGRATION
Head: Dr. Konrad Lammers

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
Hamburg Institute of International Economics
Öffentlichkeitsarbeit
Neuer Jungfernstieg 21 - 20347 Hamburg
Telefon: 040/428 34 355
Telefax: 040/428 34 451
E-mail: hwwa@hwwa.de
Internet: http://www.hwwa.de/

Gianmarco I.P. Ottaviano
Bocconi University
Istituto di Economia Politica
Via Gobbi 5 - 20136 Milan - Italy
Telefon: +39-02-58365450
Telefax: +39-051-58365314
E-Mail: gianmarco.ottaviano@uni-bocconi.it

Contents

	Page
1 INTRODUCTION	7
2 POLICY BACKGROUND: EU REGIONAL INTERVENTION	10
3 THE MODEL	15
4 THE MARKET OUTCOME	19
5 THE EFFICIENT OUTCOME	22
6 POLICY IMPLICATIONS	25
7 CONCLUSION	26
REFERENCES	27
APPENDIX	30

Figure 1: EU structural funds 11
Figure 2: EU state aid caps 13

Abstract

The global welfare implications of home market effects in trade models with imperfect competition are little understood. This paper proposes a simple model in which such implications can be easily analyzed. It shows an overall tendency of imperfectly competitive sectors to inefficiently cluster in locations that offer market access advantages. The more so the stronger the market power of firms as well as the intensity of increasing returns to scale and the lower the trade costs. As such features are likely to differ widely across sectors, those results provide theoretical ground to the promotion of regional policies that are also sector-specific and not only region-specific as currently in the EU.

J.E.L. Classification: F12, L13, R13

Keywords: economic integration, specialization, home market effect, regional disparities, regional policy

1 Introduction

Market access plays a key role in many recent models of international trade. Such models study the impact of frictions in goods and factors mobility on the location of imperfectly competitive industries in the presence of increasing returns to scale (Helpman and Krugman, 1985). Their central result is the so-called *home market* or *market size effect* (henceforth, HME), according to which, in the case of a two-country economy, the location with larger local demand succeeds in attracting a more than proportionate share of firms in the aforementioned industries. In the case of more than two countries, rather than local demand, what matters is overall market access (Krugman, 1993). For example, a small central country may have better overall market access than a large peripheral one and thus, despite its local demand disadvantage, may end up attracting a larger share of imperfectly competitive firms.[1] This pattern of demand-driven specialization maps into trade flows and generates the theoretical prediction that large central countries should be net exporters of goods produced under increasing returns and imperfect competition.[2]

From an empirical viewpoint, those predictions seem to find some support in the data. For example, Feenstra, Markusen and Rose (1998; 2001) argue that the HME is crucial to understand the empirical success of gravity equations, which explain bilateral trade flows in terms of incomes and distance between trade partners. Using disaggregated trade data from Statistics Canada World Database, they show that the HME appears to be relevant in both differentiated and homogeneous goods sectors, even though more in the former than in the latter. Using disaggregated data on UK-US trade, Weder (1997) finds that relative demand has a positive relationship with net exports as implied by the HME. Davis and Weinstein (1998; 1999) find evidence of the HME in disaggregated trade data between OECD countries. Based on disaggregated production data from Eurostat, Trionfetti (1998), Midelfart-Knarvik, Overman, Redding, and Venables (2000) as well as Brülhart and

[1] Notice that, as a consequence, with more than three countries, it is not even clear the benchmark against which to measure the local presence of imperfectly competitive firms. In other words, "more than proportionate" with respect to what?

[2] This implication derives even more strongly from 'new economic geography' models (Fujita, Krugman, and Venables, 1999), which show that, in the presence of demand and cost linkages between factors and firms, the HME can be powerful enough to cause catastrophic agglomeration of imperfectly competitive increasing-return sectors once trade costs fall below a certain threshold.

Trionfetti (2001) argue that market access is significant in explaining EU industrial specialization. Finally, analizing disaggregated industry data for Canadian and US manufacturing, also Head and Mayer (forthcoming) find supportive evidence for the HME.[3]

From a theoretical viewpoint, the focus is on the two-country case. The underpinnings of the HME are unveiled by Krugman (1980) and Helpman and Krugman (1985) with respect to imperfectly competitive industries characterized by product differentiation and free entry. Helpman (1990) qualifies previous results by stressing the demand conditions under which the HME materializes in those sectors. These conditions require the cross-elasticity between varieties of a differentiated good to be larger than the overall price-elasticity of demand for the differentiated good as a whole. Davis (1998) points out the relevance of the actual incidence of trade costs in all sectors and shows that, when transportation costs on perfectly competitive goods are considered, the HME may disappear altogether because trade cost in those sectors can effectively limit the mobility of firms in differentiated sectors. Feenstra *et al.* (1998) show that there is nothing crucial in product differentiation and free entry per se in that the HME can be expected even in homogenous-good sectors with restricted entry. All that matters is the presence of positive price-cost margins. Finally, Head, Mayer, and Ries (2000) point out that, when goods are differentiated according to their location in the geographical rather than in product space, the HME might again disappear.

To sum up, the HME seems a robust implication of trade models with imperfect competition even though its actual intensity is likely to vary from sector to sector depending on returns to scale, trade costs, entry barriers and elasticities of demand and substitution.[4] This statement, however, is subject to two main caveats. First, it holds if we focus on a sector insulated from the rest of the economy, that is, the HME is a partial equilibrium phenomenon that might be washed away by general equilibrium interactions as stressed by Helpman (1990) or Davis (1998). Second, it can be defined rigorously only with two countries that differ in nothing but size, which is not the case when technology or factor-abundance driven advantages are present or

[3] While finding supportive evidence for the HME, Head and Mayer (forthcoming) stress the overall better empirical performance of an alternative theoretical explanation grounded on country-based product differentiation (the so-called Armington model). Head, Mayer, and Ries (2000) clarify the relation between such model and the HME.

[4] See, also, Trionfetti (1998) as well as Head, Mayer, and Ries (2000) for recent surveys.

when products are differentiated according to geographical location as in the counterexample by Head, Mayer, and Ries (2000).

Nonetheless, despite its theoretical success and its promising explanatory power, the welfare properties of the HME are still little understood (see, e.g., Braunerhjelm *et al.*, 2000). In particular, the literature does not provide any answer to the fundamental question of whether the spatial distribution of economic activities implied by the HME is efficient for the economy as a whole. The aim of the present paper is to give a first answer to that question and to show how that answer can be used to discuss the desirability of current regional policies that, in most cases, aim at promoting the location of firms in peripheral regions as markets get increasingly integrated.

In particular, the paper proposes a simple two-country two-factor model with a monopolistically competitive sector. To focus on the implications of the HME and in the light of the two caveats discussed above, the monopolistically competitive market is insulated from all other markets and one country is a scaled-up version of the other. Monopolistic competitive firms are modeled à la Ottaviano, Tabuchi and Thisse (forthcoming), but differently from this paper their location is driven by footloose capital mobility rather than by workers' migration. This allows for the description of a realistic situation in which capital is freely mobile between countries while labor is not. The model reveals an overall tendency of the monopolistically competitive sector to inefficiently cluster in the country that offers better market access. The more so the stronger the market power of firms as well as the intensity of increasing returns to scale and the lower the trade costs. As these features are likely to differ widely across sectors, those results provide theoretical ground to the promotion of regional policies that are both region- and sector-specific.

The paper is in six additional parts. Section 2 describes regional state aid in the EU as a natural policy background for the theoretical analysis. Section 3 presents the model. Section 4 shows how the HME arises as a market equilibrium result. Section 5 studies the welfare properties of the market equilibrium with respect to the level of trade barriers. After pointing out that also the efficient allocation of firms exhibits a HME, it argues that the market pattern of firms' location is suboptimally biased in favor of the larger ('central') country, the more so the lower trade costs are. Section 6 shows that subsidies towards the small ('peripheral') country can be used to restore efficiency. Their amount falls as integration is deepened, but rises with the distance between center and periphery. Section 7 concludes.

2 Policy background: EU regional intervention

To define the issues at stake, a natural example is the case of the EU, which devotes staggering amounts of money to regional objectives. For istance, in 1997 the total budget of the European Union consisted of the equivalent of 87.6 billion Euro, funded mainly through VAT (42.3%), direct member states' contribution that are proportional to their GDP (40.3%), and custom duties on imports from outside the Union (16.5%). Most of the budget was devoted to two areas of intervention: 47% to the Commom Agricultural Policy (CAP) and 36.3% to Structural Funds (SFs). In relative terms, the total endowment of the SFs corresponds to 1.2% of the joint GNP of EU members states and for the period from 1994 to 1999 it reached almost 154 billion Euros (at 1994 prices). As to SFs, two main general categories of expenditures can be identified. *Economic aid* (76.4%) aims at improving the attractiveness of regions to firms both indirectly through the provision of public goods (mainly infrastructures) and directly through the compensation of individual economic enterprises. *Social aid* (9.9%) targets regional unemployment and human-capital accumulation through education and skill upgrading. Since member countries' contributions to the EU budget are proportional to their respective GDPs, the geographical allocation implies a clear pattern of *international redistribution* especially to the advantage of the countries at the EU periphery, namely Ireland, Greece and Portugal, as well as of *interregional redistribution* mainly to the advantage of the less developed peripheral regions of Italy, Spain and Eastern Germany. Such pattern is clearly visible in Figure 1, which plots the SFs coverage of population across countries as a function of an index of geographical peripherality for the planning period 2000-2006.[5]

The periphery bias of EU SFs is accompanied by the direct control over

[5] In Figure 1 the chosen measure of the peripherality of a country is the distance of its capital city from the capital of Germany (the large central country). Such measure is admittedly rough, but more sophisticated indexes would not alter the basic center-periphery pattern revealed by the picture. Population coverage refers to Objective 1 (regions suffering general underdevelopment as signalled by per-capita incomes below 75% of the EU average) plus Objective 2 (regions suffering from a concentration of declining industries as measured by observable job losses in specific sectors). The source of data is European Commission (2001a).

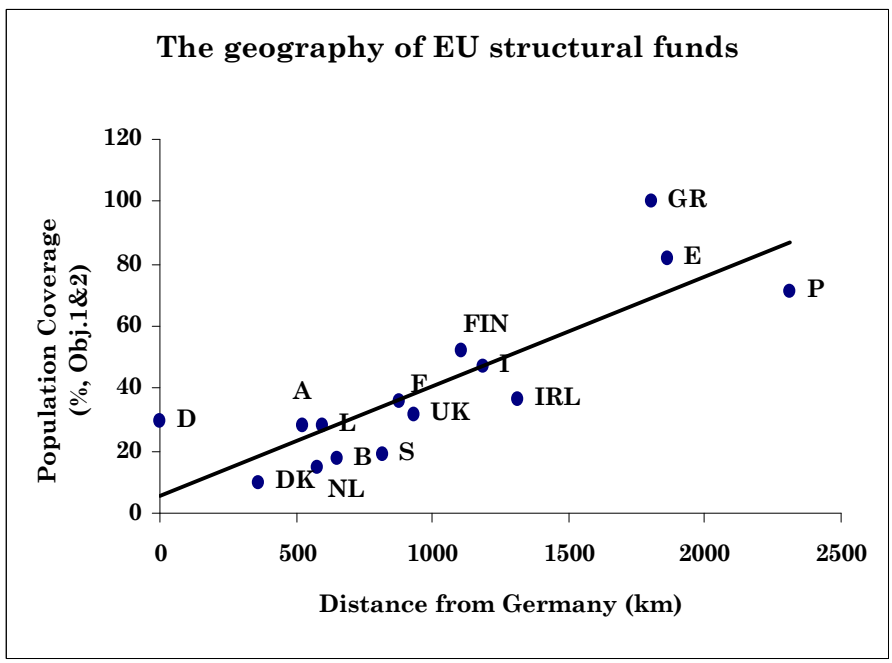

Figure 1: EU structural funds

members states' regional aid aimed at supporting productive (*initial*) investment and job creation.[6] This is achieved through *state aid caps*, that is, upper (percentage) limits to goverment support to private investments no matter whether the aid comes from local, regional, national or EU sources. As SFs state aid caps follow the logic of spatial concentration and rule out support confined to individual firms or areas of activity. The aim is to foster the development of less-favoured regions by encouraging firms to settle there and, in any case, to reduce the effects of integration on periphery-to-center relocation. The broad principle is that "[n]o trading relationship will work properly without agreed rules on the granting of subsidies" (European Commission, 1995) and it is implemented by Articles 92-94 of the Treaty of Rome (1957). In particular, as clarified by the European Commission (1995): "Article 92 specifies that state aids which distort or threaten to distort competition by favouring certain companies or the production of specific goods, and which affect trade between member states, are incompatible with the common market". Figure 2 depicts regional state aid caps across member states as a function of peripherality.[7] It matches Figure 1 in showing that, as pointed out by Braunerhjelm *et al.* (2000), "[t]he overall pattern of EU regional policy spending follows precisely the pattern that might be expected of the Commission, which is trying to achieve regional convergence [a.k.a., *cohesion*] in terms of EU GDP per capita".[8]

What has trade theory to say about the above pictures? Is the logic of redistribution from center to periphery sound as the EU faces ongoing economic integration? If so, is it about equity, efficiency, or both?

As a good approximation we can divide the possible answers into two main positions. The first is based on the neoclassical world of constant returns to scale and perfect competitition. In this world countries specialize according to their comparative advantage and each of them as a whole attains its efficient pattern of production. In other words, trade integration is Pareto improving. Nevertheless, within the same country there may be gain-

[6] Note that, generally speaking, the European Commission rules out any form of operating aid to firms.

[7] The measure of the peripherality is the same as in Figure 1. The state aid cap of a country is selected as the highest regional state aid cap within that country. The source of data is European Commission (2001b).

[8] Such pattern cannot but be enhanced by the EU prospective enlargement towards East. Indeed, the Europe agreements already allow all ten Central European and Baltic countries to define their entire territories as less-favoured regions (Objective 1).

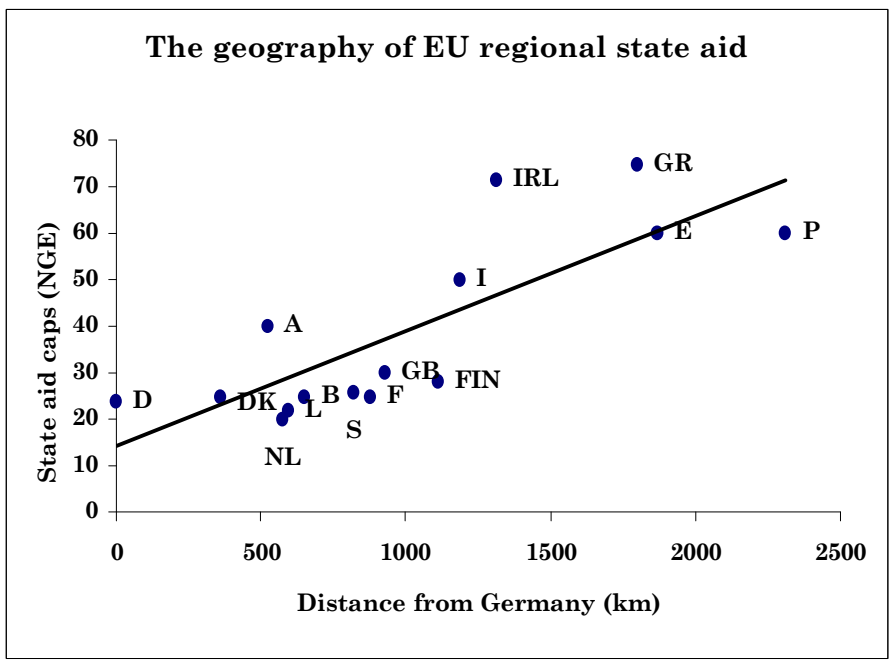

Figure 2: EU state aid caps

ers and losers. If disadvantaged interests are geographically concentrated, then there is a need for interregional transfers based on equity considerations. Notice, however, that the effects of integration (whether positive or negative) should be stronger in central regions that, by definition, are closer to international markets. At the international level, a large domestic market reduces the gains from trade integration so that, again based only on equity considerations, if anything, large countries should be compensated by small ones. Thus, the neoclassical paradigm provides little support to center-to-periphery redistribution both on equity and efficiency grounds.

The second position considers the world of increasing returns to scale and imperfect competition. As discussed in the introduction, this is the realm of the HME, that is (imperfect) trade integration makes firms relocate from peripheral to central regions. This affects negatively the former and positively the latter. The reason is that firms have market power and command rents in terms of prices set above marginal costs. These rents are extracted by firms from consumers and reduce the welfare of a country when the firms and the consumers involved belong to that same country. On the contrary, when firms and consumers are in different countries, those rents increase the welfare of the country to which firms belong. As a result, countries benefit from the expansion of their imperfectly competitive export sectors (*direct rent shifting*) as well as from an inflow of formerly foreign firms in those sectors in so far as rents are not entirely repatriated (*indirect rent shifting*) and the inflow lowers domestic consumer prices (Brander and Spencer, 1984; Helpman and Krugman, 1989; Brander, 1995). Thus, with imperfect competition, center-to-periphery relocation implies indeed that trade integration favors central countries more than peripheral ones, which may even lose. This provides equity-based support to center-to-periphery redistribution.

Compared with the foregoing results this paper moves one step further and identifies precise *sectoral* conditions under which center-to-periphery redistribution is desirable not only on equity but also on efficiency grounds. When such conditions are met, deeper trade integration enhances spatial inefficiency, thus increasing the need for redistribution. At the same time, however, it reduces the intensity of redistribution required. While these insights support the EU approach to regional intervention, at the same time they stress its incompleteness. Indeed, in so far as industry-specific characteristics are likely to determine the practical relevance of spatial inefficiency, regional policies should be not only region-specific, as they currently are in the EU, but also sector-specific.

3 The model

The analytical framework is based on the monopolitically competitive model put forth by Ottaviano and Thisse (2001) as well as Ottaviano, Tabuchi and Thisse (2002).[9] The economy consists of two countries, H and F, which are endowed with two factors, capital K and labor L. To fit the European situation, capital and labor differ in terms of international mobility. In particular, labor is assumed to be geographically immobile. Its total stock equals L, and it is evenly distributed so that a δL workers reside and work in country H. On the contrary, capital is assumed to be perfectly mobile, it is owned by workers, its total stock equals K, and it is distributed so that σK units are owned by country H residents (with $\sigma \in (0,1)$) while γK units are used in country H production (with $\gamma \in [0,1]$). Hence, $(\gamma - \sigma)K > 0$ (< 0) measures capital inflows to (outflows from) country H from (to) country F. Since the focus of the analysis is on the HME, we are interested in situations in which one 'central' country (say H) is $proportionately$ larger than the other 'peripheral' one (say F). This requires to set $\delta = \sigma > 1/2$.

In the economy there are two sectors, $modern$ and $traditional$. The modern sector is capital intensive and supplies a horizontally differentiated good under increasing returns to scale and monopolistic competition. In particular, there is an endogenous mass of firms N, each producing a single variety of the differentiated good by means of a fixed amount ϕ of capital K. The traditional sector produces a homogeneous good under constant returns to scale and perfect competition. It uses labor L as the only input with one unit of L required to produce one unit of output. This good is freely traded and is chosen as the numéraire. On the contrary, the varieties of the modern sector are traded at a cost of τ units of the numéraire per unit shipped between the two countries.

Preferences are identical across individuals and described by the following quasi-linear indirect utility function which is symmetric in all varieties:

$$V(y; p(i),\, i \in [0,N]) = -a\int_0^N p(i)di + \frac{b+cN}{2}\int_0^N [p(i)]^2 di \qquad (1)$$

[9] The model adopted reproduces the basic features of Dixit and Stiglitz (1977) using different functional forms. As discussed in Ottaviano, Tabuchi, and Thisse (2002), compared with the standard CES implementation of Dixit and Stiglitz's insights, the present model comparative advantage lies in neater comparative statics results and more straightfoward welfare analysis with heterogenous agents.

$$-\frac{c}{2}\left[\int_0^N p(i)di\right]^2 + y + \overline{q}_0$$

where $p(i)$ is the price of variety $i \in [0, N]$, y the consumer's income, and $\overline{q}_0$ her initial endowment of the numéraire. In (1), $a > 0$ expresses the intensity of preferences for the differentiated product with respect to the numéraire; $b > 0$ means that the representative consumer is biased toward a dispersed consumption of varieties, thus reflecting a love for variety; $c > 0$ expresses the substitutability between varieties so that the higher c, the closer substitutes the varieties. Finally, the initial endowment $\overline{q}_0$ in the numéraire is assumed to be large enough for the consumption of the numéraire to be strictly positive at the market equilibrium and optimal solutions.

Labor market clearing implies that the number n_H of firms belonging to the modern sector and located in country H is equal to:

$$n_H = \gamma K/\phi \tag{2}$$

so that the number of firms in F is

$$n_F = (1-\gamma)K/\phi \tag{3}$$

Consequently, the total number of firms (varieties) in the economy is fixed by endowments and technology and equal to $N = K/\phi$.

Entry and exit are free so that profits are zero in equilibrium. Hence, (2) and (3) imply that any change in the number of firms located in one country originates from a corresponding change in the locally employed stock of capital. By (2) and (3), the demand and supply of capital in each country are equal. As a result, the corresponding equilibrium returns to capital are determined by a bidding process among firms which ends when no firm can earn a strictly positive profit at the equilibrium market prices.

Firms are assumed to take advantage of positive trade costs to segment markets, that is, each firm sets a price specific to the market in which its product is sold. This assumption follows from empirical work showing that, even within a unified economic area, firms succeed to price discriminate between spatially separated markets (McCallum, 1995; Head and Mayer, 2000). As shown below, in equilibrium arbitrage is not profitable to third parties.

In what follows, we focus on country H. Things pertaining to country F can be derived by symmetry. Using the assumption of symmetry between

varieties and Roy's identity, individual demands for a representative firm in H are given by:

$$q_{HH} = a - (b + cN)\, p_{HH} + cP_H \tag{4}$$

and

$$q_{HF} = a - (b + cN)\, p_{HF} + cP_F \tag{5}$$

where p_{HH} (p_{HF}) is the price set in H (F) by a firm located in H and

$$P_H \equiv n_H\, p_{HH} + n_F\, p_{FH}$$

$$P_F \equiv n_H\, p_{HF} + n_F\, p_{FF}$$

Clearly, P_H/N and P_F/N can be interpreted as the price indices prevailing in countries H and F.

A representative firm in H maximizes its profits, which, after using (4) and (5), are defined by:

$$\begin{aligned} \Pi_H \;=\; & p_{HH}\, [a - (b + cN)\, p_{HH} + cP_H]\, \sigma L + \\ & (p_{HF} - \tau)\, [a - (b + cN)\, p_{HF} + cP_F]\, (1 - \sigma) L - \phi r_H \end{aligned} \tag{6}$$

where r_H is the return to capital prevailing in H.

Market prices are obtained by maximizing profits while capital returns are determined as described above by equating the resulting profits to zero. Since we have a continuum of firms, each one is negligible in the sense that its action has no impact on the market. Hence, when choosing its prices, a firm in H accurately neglects the impact of its decision over the two price indices P_H and P_F. In addition, because firms sell differentiated varieties, each one has some monopoly power in that it faces a demand function with finite elasticity. On the other hand, since the price index enters the demand function as an additive term (see (4) and (5)), a firm must account for the distribution of the firms' prices through some aggregate statistics, given here by the average market price, in order to find its equilibrium price. As a consequence, the market solution is given by a Nash equilibrium with a continuum of players in which prices are interdependent: each firm neglects its impact on the market but is aware that the market as a whole has a non-negligible impact on its behavior.

Since profit functions are concave in own price, solving the first order conditions for profit maximization with respect to prices yields the equilibrium prices:

$$p_{HH} = \frac{1}{2} \frac{2a + \tau cN(1 - \gamma)}{2b + cN} \tag{7}$$

$$p_{FF} = \frac{1}{2}\frac{2a + \tau cN\gamma}{2b + cN} \tag{8}$$

$$p_{HF} = p_{FF} + \frac{\tau}{2} \tag{9}$$

$$p_{FH} = p_{HH} + \frac{\tau}{2} \tag{10}$$

which depend on the total number of active firms as well as on their distribution between the two countries.

Substracting τ from (9) and (10), we see that firms' prices net of trade costs are positive regardless of their spatial distribution if and only if

$$\tau < \tau_{trade} \equiv \frac{2a\phi}{2b\phi + cK} \tag{11}$$

The same condition must hold for consumers in F (H) to buy from firms in H (F), i.e. for the demand (5) evaluated at the prices (7) and (8) to be positive for all γ. From now on, condition (11) is assumed to hold. Consequently, we consider a setting in which there is a priori intra-industry trade.

Using (11) we observe that more firms in the economy lead to lower market prices for the same spatial distribution $(\gamma, 1-\gamma)$ because there is more competition in each local market. Similarly, both the prices charged by local and foreign firms fall when the mass of local firms increases because competition is fiercer. Equilibrium prices also rise when the degree of product differentiation, inversely measured by c, increases provided that (11) holds. Moreover, it can be easily checked that $p_{HF} - p_{HH} < \tau$ (i.e., there is $dumping$) so that the prohibition of arbitrage associated with the assumption of segmented markets is not binding.

Finally, local sales rise with τ because of the higher protection enjoyed by the local firms but exports fall for the same reason. By using (6), (7), and (9), it is easy to check that the equilibrium operating profits earned by a firm established in H on each separated market are as follows:

$$\Pi_{HH} = (b + cN)\, p_{HH}^2\, \sigma L$$

where Π_{HH} denotes the profits earned in H while the profits made from selling in F are

$$\Pi_{HF} = (b + cN)\, (p_{HF} - \tau)^2\, (1 - \sigma) L$$

Thus, an increase in the number of firms in one country decreases the operating profits of local sales due to tougher local competition: the equilibrium price falls as well as the quantity of each variety bought by each consumer.

The individual consumer surplus S_H in country H associated with the equilibrium prices (7) and (10) is then as follows (a symmetric expression holds in country F):

$$\begin{aligned} S_H(\gamma) &= -a[\gamma\, p_{HH} + (1-\gamma)\, p_{FH}]N \\ &\quad + \frac{b+cN}{2}\,[\gamma\, p_{HH}^2 + (1-\gamma)\, p_{FH}^2]\, N - \frac{c}{2}[\gamma\, p_{HH} + (1-\gamma)\, p_{FH}]^2 N^2 \end{aligned} \tag{12}$$

which can be shown to be quadratic in γ. Differentiating twice this expression with respect to γ shows that $S_H(\gamma)$ is concave. Furthermore, (11) implies that $S_H(\gamma)$ is always increasing in γ over the interval $[0,1]$. Hence, as more firms enter in H, the surplus of residents rises because local competition becomes fiercer; however, this effect gets weaker and weaker as the number of local firms increases.

The equilibrium return to capital prevailing in country H can be obtained by evaluating $r_H(\gamma) = (\Pi_{HH} + \Pi_{HF})/\phi$, thus yielding the following expression:

$$\begin{aligned} r_H(\gamma) &= \frac{b\phi + cK}{4(2b\phi + cK)^2\phi^2}\Big\{[2a\phi + \tau cK(1-\gamma)]^2\,\sigma L \\ &\quad + [2a\phi - 2\tau b\phi - \tau cK(1-\gamma)]^2\,(1-\sigma)L\Big\} \end{aligned} \tag{13}$$

which is also quadratic in γ. Standard, but cumbersome, investigations reveal that $r_H(\gamma)$ is convex and decreasing in γ. In other words, the equilibrium return to capital wage falls with the local number of firms so that, while $S_H(\gamma)$ rises, $r_H(\gamma)$ decreases with γ. This effect gets weaker and weaker as the number of local firms increases because the larger their number, the weaker the marginal impact of a new entrant on the intensity of local competition. Moreover, inspection of the square bracketed terms reveals that operating profits per unit sold are larger on domestic than on distant sales and the more so the smaller the fraction of domestic firms.

4 The market outcome

We are now ready to determine the equilibrium location of firms as the result of the international allocation of capital. Since it is capital flows that

determine the location of firms, an equilibrium arises when no capital owner can earn strictly higher returns by changing the country serviced by her capital endowment. This happens for $0 < \gamma < 1$ whenever capital returns are equalized in the two countries:

$$r_H(\gamma) = r_F(\gamma) \tag{14}$$

and for $\gamma = 1$ [$\gamma = 0$] whenever $r_H(1) \geq r_F(1)$ [$r_F(0) \geq r_H(0)$].[10] In these latter cases the modern sector is clustered in one country only, with the other country completely specialized in the production of the traditional good.

Using (13) as well as the corresponding expression for country F in (14), the differential return on capital can be expressed as:

$$r_H(\gamma) - r_F(\gamma) = C\left\{\tau[4a\phi - \tau(2b\phi + cK)](\sigma - 1/2) - \tau^2 cK(\gamma - \sigma)\right\} \tag{15}$$

with $C \equiv [(b\phi + cK)K]/[2\phi^2(2b\phi + cK)] > 0$. The differential is, thus, a decreasing linear function of γ.

The right hand side of (15) shows that the equilibrium spatial allocation of capital is determined by the interaction of the two terms inside the curly bracket. The first term depends on the spatial distribution of consumers (σ) while it is independent from the location of firms (γ). Since the coefficient of $(\sigma - 1/2)$ is positive in so far as (11) holds, that term measures a $market\ access$ advantage due to trade costs saving: were the overall spatial distribution of firms to mirror the distribution of consumers ($\gamma = \sigma$), it would nonetheless be better to be located in the larger country because, as mentioned above, operating profits per unit sold are larger on domestic than on distant sales. This is not necessarily true when there is a more than proportionate presence of firms in the larger country ($\gamma > \sigma$). In that case, the second term in (15) points out that there is a $market\ crowding$ penalty, which derives from the fact that, as the fraction of firms in the larger market grows, operating profits per unit sold fall on domestic sales and rise on distant ones (see (13)). This increases the incentive to export and the associated trade cost burden. Indeed, if no country offered better market access than the other ($\sigma = 1/2$), then the operating profits maximizing allocation of firms would mirror the spatial disitribution of consumers ($\gamma = \sigma$) as that would minimize trade costs.

[10] Since $r_H(\gamma)$ is decreasing while $r_F(\gamma)$ is increasing in γ, if they cross, they do so only once.

Equation (15) also reveals that the balance is tilted in favor of market access when a and ϕ are large (given (11)) as well as when b, c, and K are small. Under such circumstances, the elasticities of demand and substitution of a typical variety (see (4) and (5)) are both small, thus implying that a large component of operating profits is independent from the overall distribution of firms. In particular, as intuition would have it, in the limit case of monopoly ($c = 0$) only market access considerations matter since a firm's operating profits are unrelated to other firms' locations. Finally, the balance between market access and crowding is also affected by the level of trade costs in that access considerations dominate for low trade costs, while crowding concerns are crucial for large trade costs. The reason why is that, with lower trade costs, a larger fraction of operating profits is independent from the overall location of firms.

Solving (13) for γ, we obtain the equilibrium location of firms:

$$\gamma^M = \sigma + \frac{4a\phi - \tau(2b\phi + cK)}{\tau cK}(2\sigma - 1) \tag{16}$$

so that γ^M is alway larger than σ whenever $\sigma > 1/2$ and less than 1 whenever τ is larger than

$$\tau_{cluster} \equiv \frac{4a\phi(2\sigma - 1)}{2b\phi(2\sigma - 1) + cK} \tag{17}$$

When τ falls short of this threshold the modern sector is clustered inside country H and country F is completely specialized in the production of the traditional good. Therefore, the incomplete specialization of F is compatible with international trade flows only if $\tau_{trade} > \tau_{cluster}$ i.e.

$$\sigma < \frac{1}{2} + \frac{cK}{4(b\phi + cK)} \tag{18}$$

which shows that the modern sector is more likely to cluster the larger country H (larger σ), the higher the degree of product differentiation (lower c), the more intense the returns to scale (larger ϕ). When (18) is violated trade always leads to complete specialization of the larger country in the production of the modern good.

The fact that γ^M is always larger than σ (given $\sigma > 1/2$) reveals the existence of a HME: the larger country H attracts a more than proportionate number of modern firms. In particular, we have:

$$\frac{d\gamma^M}{d\sigma} = \frac{2\phi(2a - \tau b)}{\tau cK} \tag{19}$$

which is larger than 1 whenever τ is smaller than:

$$\tau_{HME} \equiv \frac{4a\phi}{2b\phi + cK} \tag{20}$$

This is indeed the value of τ that erases the market access advantage in (15).

Therefore, since $\tau_{HME} = 2\tau_{trade} > \tau_{trade}$, *trade is always associated with a HME*. Moreover, by comparing (20) with (17), it is easy to notice that τ_{HME} is the maximum value of $\tau_{cluster}$ achieved at $\sigma = 1$. Therefore, we have also $\tau_{HME} > \tau_{cluster}$.

The equilibrium capital flow from F to H is:

$$CF_{FH} \equiv (\gamma^M - \sigma)K = \frac{(2\sigma - 1)[4a\phi - \tau(2b\phi + cK)]}{2\tau c} \tag{21}$$

which is positive given (11). Of course, this is also a measure of relocation defined as $(\gamma^M - \sigma)N = CF_{FH}/\phi$. By (21), the extent of relocation is a decreasing function of c as well as K and an increasing function of a as well as ϕ for the same reasons discussed above. Moreover, it is also a decreasing function of τ, implying that capital flows grow as trade costs fall. Interestingly enough, since, as it can be easily shown, the total trade volume is also a decreasing function of τ, the model predicts that foreign direct investment and international trade grow together as economic integration deepens. As a result, the larger country H increasingly exchanges the modern good against the services of both factors. In the case of capital the inflow is direct, while in the case of labor it is embodied in traditional imports.

5 The efficient outcome

In principle the model has two potential sources of inefficiency. On the one side, for a given spatial distribution γ, when pricing above marginal cost, firms do not take into account the social loss in terms of consumer surplus. On the other, for given prices, when choosing location they do not consider the impact of their decisions on competitors' profits and consumers' surpluses. Notice however that, differently from Dixit and Stiglitz (1977), in the present setting the total number of firms N is always efficient since, as a consequence of (2) and (3), that number is determined by the total endowment of K and the technology parameter ϕ.

Consider initially a *first-best planner* who has enough instruments to eliminate all sources of inefficiency. In other words, assume that the planner is able (i) to assign any number of firms to a specific country and (ii) to use lump-sum transfers from workers to pay for the loss firms may incur while pricing at marginal cost. The planner chooses γ in order to maximize the following social welfare function (recall that individual utilities are quasi-linear):

$$\begin{aligned} W(\gamma) &= S_H(\gamma)\,\sigma L + S_F(\gamma)\,(1-\sigma)L \\ &\quad + r_H(\gamma)\,\gamma K + r_F(\gamma)\,(1-\gamma)K + \text{constant} \end{aligned} \tag{22}$$

which is simply the sum of all workers' indirect utilities and where all prices have been set equal to marginal cost:

$$p_{HH} = p_{FF} = 0 \qquad \text{and} \qquad p_{HF} = p_{FH} = \tau$$

The planner actually engages in international trade as long as consumers want to buy foreign varieties (q_{HF}, $q_{FH} > 0$), which is the case if and only if trade costs are such that:

$$\tau < \tau^O_{trade} \equiv \frac{a\phi}{b\phi + cK} \tag{23}$$

Since $\tau^O_{trade} < \tau_{trade}$, as in Brander and Krugman (1983), there is a range of relatively large but not prohibitive trade costs in which inefficient commerce takes place at the market outcome due to the dumping behavior of firms.

Up to a positive multiplicative constant, the first order condition of the planner's problem is:

$$\tau[2a\phi - \tau(b\phi + cK)](\sigma - 1/2) - \tau^2 cK(\gamma - \sigma) = 0 \tag{24}$$

Straightforward comparison with (15) shows that the planner gives less weight to market access than the decentralized outcome. This happens for the following reason. Starting with $\gamma = \sigma > 1/2$, as γ rises, the local sales of each firm in H go down while its distant sales go up. The opposite happens to firms in F. However, given $\gamma = \sigma > 1/2$, the net result is an increase in aggregate shipments and thus in aggregate trade costs. The fact that the planner internalizes this effect explains why the social weight of market access falls short of the private one.[11]

[11] The Appendix solves the trade costs minimization problem to argue that transport saving is indeed the driving force behind the HME. That is why some degree of HME is always optimal.

Solving (24) in γ yields the first best spatial distribution of firms:

$$\gamma^O \equiv \sigma + \frac{2a\phi - \tau(b\phi + cK)}{\tau cK}(2\sigma - 1) > \sigma \tag{25}$$

which, under (23), implies that also the planner delivers a HME by allocating a more than proportionate share of firms in the larger country. However, since clearly $\gamma^O < \gamma^M$, the market outcome has an inefficiently large number of firms in country H. Moreover, such inefficiency is larger the stronger the decentralized HME: $d\gamma^O/d\sigma = (1/2)(d\gamma^M/d\sigma) > 1$.

While the first-best planner is an interesting benchmark, in reality marginal cost pricing is diffcult to implement because of the lack of the lump-sum instruments involved. It is therefore useful to consider the choice of a *second-best planner* who is able to assign any number of modern firms to a specific country but is unable to use lump-sum transfers from workers to firms. In this case, the planner chooses γ in order to maximize (22) evaluated at market prices (7)-(10) so that she engages in trade if and only if (11) holds.

The second-best first order condition is:

$$\tau[8a\phi(3b\phi+cK)-3\tau(2b\phi+cK)^2](\sigma-1/2)-\tau^2cK(8b\phi+3cK)(\gamma-\sigma) = 0 \tag{26}$$

which implies that, relative to market crowding, market access is overweighed with respect to the first best, and underweighed with respect to the decentralized outcome. As the first best planner, the second best one internalizes the adverse trade cost surcharge that additional firms in H impose on local incumbents. However, since under monopolistically competitive pricing firms absorb part of the trade costs ($p_{HF} - p_{HH} < \tau$), the surcharge is smaller than under marginal cost pricing.

Solving (26) in γ yields the second-best spatial distribution of firms:

$$\gamma^S \equiv \sigma + \frac{8a\phi(3b\phi + cK) - 3\tau(2b\phi + cK)^2}{2\tau cK(8b\phi + 3cK)}(2\sigma - 1) \tag{27}$$

which, given (11), implies that also the second-best planner delivers a HME. In particular, we have $d\gamma^S/d\sigma = [(6b\phi+2cK)/(8b\phi+3cK)](d\gamma^M/d\sigma)$ so that $d\gamma^M/d\sigma > d\gamma^S/d\sigma > d\gamma^O/d\sigma > 1$.

By simple inspection $\gamma^S < \gamma^M$ with

$$\gamma^M - \gamma^S = \frac{\phi(2a - \tau b)(2b\phi + cK)(2\sigma - 1)}{\tau cK(8b\phi + 3cK)} > 0 \tag{28}$$

so that the market outcome leads to too much concentration also with respect to the second-best allocation; the more so the stronger the HME. It is also readily verified that $\gamma^O < \gamma^S$ with:

$$\gamma^S - \gamma^O = \frac{\phi(2a - \tau b)(4b\phi + cK)(2\sigma - 1)}{2\tau cK(8b\phi + 3cK)} > 0 \tag{29}$$

This suggests that both sources of inefficiency work in the same direction by supporting a spatial distribution of firms that is too uneven. Notice also that $(\gamma^S - \gamma^O) < (\gamma^M - \gamma^S)$ so that the second-best is closer to the first-best than to the market outcome. Finally, the discrepancies between γ^M, γ^S, and γ^O grow as τ falls: *economic integration widens the gap between market and efficient outcomes*.

6 Policy implications

How can a policy maker implement the efficient spatial distribution of firms? Are EU policies going in the right direction? To answer these questions we have to consider once more the logic of the model. For concreteness, we target the second-best allocation.

Modern firms' location is determined by capital owners decisions on the provision of their services. These decisions are guided by the differential between the returns to capital in the two countries, $[r_H(\gamma^S) - r_F(\gamma^S)]$. Since such differential decreases with γ (see (13)) being zero at γ^M and γ^S is smaller than γ^M, then it must be that $[r_H(\gamma^S) - r_F(\gamma^S)] > 0$, that is, at γ^S there is a positive gap between capital returns in the larger country H and in the smaller country F. Any policy tool that is able to fill in that gap will achieve the second best.

The related gain in terms of overall welfare will be:

$$W(\gamma^S) - W(\gamma^M) = \frac{(2\sigma - 1)^2(2a - \tau b)^2(b\phi + cK)L}{8c(8b\phi + 3cK)} \tag{30}$$

which increases as the level of trade costs falls. In other words, *economic integration increases the welfare loss due to the inefficient spatial disitribution of firms at the market outcome*.

With the practices of the EU in mind, we consider an investment subsidy to the fixed costs of firms in F levied through income taxation.[12] Let s^* be

[12] Given quasi-linear utility and workers' ownership of capital, the exact way income taxes are raised is immaterial in so far as we think in terms of a net investment subsidy.

the *optimal investment subsidy* per unit of capital invested in F. Then, s^* is such that $[r_H(\gamma^S) - r_F(\gamma^S) - s^*] = 0$ implying:

$$s^* = \frac{\tau L(2a - \tau b)(b\phi + cK)(2\sigma - 1)}{2\phi(8b\phi + 3cK)} \tag{31}$$

This shows that the optimal subsidy is an increasing function of the trade costs τ (since (11) holds), the total capital stock K, the substitutability between varieties c, while it is a decreasing function of the intensity of returns to scale ϕ. Consequently, as trade costs fall, the optimal subsidy shrinks: *as countries get more integrated,* the overall welfare loss due to the inefficient distribution of modern firms rises, but, at the same time, *the amount of international redistribution needed to restore efficiency falls*. The explanation is that, as trade cost fall, firms become increasingly footloose. While, on the one side, this fosters their inefficient concentration in the larger country, on the other it makes them more sensitive to any differential in subsidies.[13] The same effects are associated with falling c and rising ϕ.

7 Conclusion

A distinguishing feature of new trade models is the so-called home market effect, that is, in the case of two countries, the more than proportionate location of imperfectly competitive increasing-return sectors in the larger country. However, despite its centrality and the distortions at its origin, so far the home market effect has attracted surprisingly little attention in terms of its global welfare implications.

This paper represents a first step in the direction of filling that gap. Using a simple new trade model, it has shown that sectors characterized by increasing returns to scale and imperfect competition tend to be inefficiently over-concentrated in the larger country, the more so the lower trade costs are. This implies that, in the process of integration, the demand for active policy intervention to reduce economic disparities between large ('central')

[13] One may wonder then why the EU regional budget has been growing through time? The answer is ongoing enlargement. In the present model, the simplest way to capture the joint phenomena of deeper integration between old central members and additional inclusion of new peripheral ones is to have σ rising (the center grows) and τ rising (the periphery gets more distant) at the same time. On both counts, by (31), the optimal subsidy rises.

countries and small ('peripheral') ones may stand not only on equity but also on efficiency grounds. However, industry-specific characteristics are likely to determine the practical relevance of spatial inefficiency. Accordingly, differently from current EU practice, regional policies should be targeted not only to specific peripheral regions but also to specific sectors characterized by steeply increasing returns to scale and strong firms' market power.

Stemming from a first attempt, these results are obviously preliminary and should be qualified by studying more general models of the HME. As discussed in the introduction, in the wake of existing results, such models should be built on two main pillars. First, the HME is essentially a partial equilibrium phenomenon that might be washed away by general equilibrium interactions. Consequently, one should focus on a sector insulated from the rest of the economy. Second, the HME can be defined only when countries differ in nothing but market access. So far, in the absence of a benchmark measure of market access with many countries, one should concentrate on a two-country economy where market access is simply captured by local market size.

References

[1] Brander, J.A. (1995) Strategic trade policy. In Grossman, Gene M. and Kenneth Rogoff (eds.) *Handbook of International Economics*, Vol.III (Amsterdam: North Holland).

[2] Brander, J.A. and P.R. Krugman (1983) A 'reciprocal dumping' model of international trade, *Journal of International Economics* 15, 313-323.

[3] Brander, J.A., and B.J. Spencer (1984) Tariff protection and imperfect competition. In H. Kierzkowski (ed.) *Monopolistic Competition and International Trade* (Oxford: Oxford University Press).

[4] Braunerhjelm, P., R. Faini, V. Norman, F. Ruane, and P. Seabright (2000) *Integration and Regions in Europe: How the Right Policies Can Prevent Polarization*, Monitoring European Integration 10 (London: Centre for Economic Policy Research).

[5] Brülhart, M. and F. Trionfetti (2001) Industrial structure and public procurement: theory and empirical evidence, *Journal of Economic Integration* 16, 106-127.

[6] Davis D.R. (1998) The home market effect, trade, and industrial structure, *American Economic Review* 88, 1264-1276.

[7] Davis D.R. and D. E. Weinstein (1998) Market access, economic geography and comparative advantage: an empirical assessment, NBER Working Paper No.6787.

[8] Davis D.R. and D. E. Weinstein (1999) Economic geography and regional production structure: an empirical investigation, *European Economic Review* 43, 379-407.

[9] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *American Economic Review* 67, 297-308.

[10] European Commission (1995) Free movement of people, DGX *European Dialogue* 1.

[11] European Commission (1996) *First Report on Economic and Social Cohesion* (Luxembourg: Office for Official Publications of the European Communities).

[12] European Commission (2001a), *Regional State Aid Map 1999*, DG Competition, http://europa.eu.int/comm/competition.

[13] European Commision (2001b), *EU Structural Funds 2000-2006*, DG Employment and Social Affairs, http://europa.eu.int/comm/employment_social.

[14] Feenstra R.C., J.R. Markusen and A.K. Rose (1998) Understanding the home market effect and the gravity equation: the role of differentiated goods, CEPR Discussion Paper No.2035.

[15] Feenstra R.C., J.R. Markusen and A.K. Rose (2001) Using the gravity equation to differentiate between alternative models of trade, *Canadian Journal of Economics*.34.

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

[17] Head, K. and T. Mayer (2000) Non-Europe. The magnitude and causes of market fragmentation in the EU, *Weltwirtschaftliches Archiv* 136, 285-314.

[18] Head, K. and J. Ries (forthcoming) Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade, *American Economic Review*.

[19] Head, K., T. Mayer, and J. Ries (2000) On the pervasiveness of the home market effect, University of British Columbia, mimeo.

[20] Helpman E. (1990) Monopolistic competition in trade theory, *Special Paper in International Economics* 16, Princeton University, International Finance Section.

[21] Helpman E. and P.R. Krugman (1985) *Market Structure and Foreign Trade* (Cambridge (Mass.): MIT Press).

[22] Helpman E. and P.R. Krugman (1989) *Trade Policy and Market Structure* (Cambridge (Mass.): MIT Press).

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

[24] Krugman, P.R. (1993) The hub effect: or, threeness in international trade. In Ethier, W.J., E. Helpman, and J.P. Neary (eds.) *Theory, Policy and Dynamics in International Trade* (Cambridge: Cambridge University Press).

[25] McCallum, J. (1995) National borders matter: Canada-US regional trade patterns, *American Economic Review* 85, 615-623.

[26] Midelfart-Knarvik, K.H., H.G. Overman, S.J. Redding, and A.J. Venables (2000) The location of European Industry, European Commission, DG Economic and Financial Affairs, Economic Paper No.142.

[27] Ottaviano, G.I.P. and J.-F. Thisse (2001). Integration, agglomeration and the political economics of factor mobility, *Journal of Public Economics*, forthcoming.

[28] Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse (2002). Agglomeration and trade revisited, *International Economic Review*, forthcoming.

[29] Trionfetti F. (1998) On the home market effect: theory and empirical evidence, CEP Working Paper No.987, London School of Economics.

[30] Weder, R. (1997) British and American exports: an empirical analysis of comparative home-market advantage, University of British Columbia, mimeo.

Appendix

The aim of this appendix is to show that the essence of the reason why the first best planner delivers a HME is trade costs minimization.[14]. Accordingly, consider the problem of minimizing aggregate trade costs under marginal cost pricing ($p_{HH} = p_{FF} = 0$, $p_{HF} = p_{FH} = \tau$):

$$\min_{\gamma} \tau[q_{FH}(1-\gamma)\sigma + q_{HF}\gamma(1-\sigma)] \tag{A1}$$

where, given (4) and (5), quantities shipped are:

$$q_{FH} = a-(b+cN)p_{FH}+cN[\gamma p_{HH}+(1-\gamma)p_{FH}] = a-(b+cN)\tau+\tau c(1-\gamma)N \tag{A2}$$

$$q_{HF} = a-(b+cN)p_{HF}+cN[\gamma p_{HF}+(1-\gamma)p_{FF}] = a-(b+cN)\tau+\tau c\gamma N \tag{A3}$$

Thus, after substitution of (A2) and (A3), (A1) can be rewritten as:

$$\min_{\gamma} \tau\left\{[a-\tau(b+cN)][(1-\gamma)\sigma+\gamma(1-\sigma)] + \tau cN[(1-\gamma)^2\sigma+\gamma^2(1-\sigma)]\right\}$$

where the first term inside the curly brackets refers to the spatial distribution of the component of individual import demands that is common to all consumers no matter where they reside. The associated trade costs are clearly minimized when all firms are located in the bigger region (whenever trade costs are low enough to allow for trade). As to the second term inside the curly brackets, it concerns the spatial distribution of the component of individual import demands that depends on the location of firms. This is a convex function of γ with a minimum at $\gamma = \sigma$.

The corresponding necessary condition for minimization is:

$$\tau[a\phi-\tau(b\phi+cK)](\sigma-1/2) - \tau^2 cK(\gamma-\sigma) = 0 \tag{A4}$$

[14] Indeed, in the present setting, since marginal costs are zero, trade costs minimization is equivalent to total costs minimization.

where we have used the fact that $N = K/\phi$. Condition (A4) shows that trade costs minimization gives less weight to market access than the first best outcome. It can be readily solved in γ to yield

$$\gamma^T = \sigma + \frac{[a\phi - \tau(b\phi + cK)]}{2\tau cK}(2\sigma - 1) > \sigma \tag{A5}$$

where the inequality is granted by (23).

Expression (A5) implies the existence of a HME:

$$\frac{d\gamma^T}{d\sigma} = \frac{\phi(a - \tau b)}{\tau cK} > 1$$

which is, however, less pronounced than in the first best case ($\gamma^O > \gamma^T$) as it can be readily verified by comparing (A5) with (25).

Given previous discussions, the HME is entirely driven by the minimization of trade costs associated with the component of individual import demands that is common to all consumers no matter where they reside.