

**National Disparities
and
Regional Allocation of Resources:
A Positive Framework***

Gianmarco I.P. Ottaviano**

University of Bologna and CEPR

THIRD DRAFT

9 September 2005

Abstract

Deeper integration of MERCOSUR aims at reducing the barriers to the international mobility of goods, factors and ideas as well as at promoting effective policy coordination among member countries. Lower barriers make the interactions among customers and suppliers in the integrated area increasingly tight, thus fostering the creation of a common market within the MERCOSUR economic space. The present paper proposes a theoretical framework to assess the economic impact and the welfare implications of the resulting reallocation of resources across firms and countries. Its final purpose is to answer the central questions whether and how the distribution of the associated costs and benefits might create resistance from national interests to move towards deeper integration.

* Paper prepared within the framework of “Deeper Integration of MERCOSUR: Dealing with Disparities”, a joint initiative between the Integration and Regional Programs Department and the Operations Department 1 of the Inter-American Development Bank. The author is grateful to Mauricio Moreira and Christian Volpe Martincus for advise on MERCOSUR stylized facts. He is also indebted with seminar participants at the Inter-American Development Bank in Washington DC, those at the conference “Deepening Integration of MERCOSUR: Dealing with Disparities” in Rio de Janeiro as well as with the IDB-INT/RE1 team for helpful comments.

** Dipartimento di Scienze Economiche, Università di Bologna, Strada Maggiore 45, 40125 Bologna, Italy. E-mail: ottavian@economia.unibo.it.

.1. Introduction

Deeper integration of MERCOSUR aims at reducing the barriers to the international mobility of goods, factors and ideas as well as at promoting effective policy coordination among member countries. Lower barriers make the interactions among customers and suppliers in the integrated area increasingly tight, thus fostering the creation of a common market within the MERCOSUR economic space. The present paper proposes a theoretical framework to assess the economic impact and the welfare implications of the resulting reallocation of resources both between and within countries. Its final purpose is to answer the central questions whether and how the distribution of the associated costs and benefits might create resistance from national interests to move towards deeper integration. In particular, as argued in Section .2, one of the highest obstacles to deeper integration is increasingly represented by asymmetries between member countries and the associated discontent of smaller countries about the distribution of related gains and losses.

The foregoing issues are tackled from the specific point of view of ‘new economic geography’ (henceforth, simply NEG), an approach to economic geography firmly grounded on recent developments in mainstream industrial organization and international trade theory. After more than a decade since the seminal work by Krugman (1991), NEG has grown into a mature body of literature as testified by a rich list of surveys and textbooks such as Ottaviano and Puga (1998), Fujita et al (1999), Neary (2001), Ottaviano and Thisse (2001, 2004), Fujita and Thisse (2002), Baldwin et al (2003), Ottaviano (2003), Ottaviano and Pinelli (2004).

Based on these two last surveys, Section .3 presents the building blocks and the main insights of this literature with particular reference to the interaction among firms in imperfectly competitive markets. First, *ceteris paribus*, product market competition promotes the geographical dispersion of economic activities. Second, international differences in production costs foster agglomeration in low-cost countries. Third, international differences in local market size foster agglomeration in larger countries. Fourth, productivity differences across firms foster dispersion. Fifth, trade barriers have a non-linear impact on the balance between agglomeration and dispersion forces. Last, agglomeration forces may generate ‘cumulative causation’ among firms’ location decisions.

NEG models, however, typically neglect some important micro-economic effects of trade integration. Indeed, Section .2.3 argues that, quite surprisingly, MERCOSUR *per se* seems to have had little impact on the aggregate economic performance of member countries. The reason is the parallel implementation of other important policy reforms involving both labour and financial markets. The separate impact of trade liberalization, instead, can be detected at the finely disaggregated level. For instance, recent empirical investigations on Argentina and Uruguay reveal

that policy reforms have caused intense capital and labour reallocation that have happened within rather than between sectors. That is, most action has involved the contraction of less productive establishments and the expansion of more productive ones within the very same industries. In this respect, among all policy reforms, the largest impact comes from trade liberalization rather than from financial and labour reforms.

To illustrate the theoretically underpinnings of such microeconomic dynamics, Section .4 enriches the analytical framework of NEG in the wake of Melitz (2003), Melitz and Ottaviano (2005) as well as Bernard, Redding and Schott (2004). The result is a rich set of predictions on the effects of trade liberalization in terms of industry performance measures. These highlight the role of available technologies ('cost-saving attraction'), local market size ('market-seeking attraction'), and access to other markets ('accessibility'). Ceteris paribus, as exemplified in Section .5, the enriched analytical framework produces the following implications:

- Due to market-seeking attraction, larger local markets are characterized by tougher competition. This generates richer product variety, higher productivity, lower prices, and higher welfare.
- Due to cost-saving attraction, in absolute terms technologically advanced regions are characterized by tougher competition. In relative terms, regions feature tougher competition in their sectors of comparative advantage. This generates richer product variety, higher productivity, lower prices, and higher welfare.
- Due to accessibility, 'hubs' (i.e. regions that occupy a central place within the trade network) are characterized by tougher competition and, therefore, richer product variety, higher productivity, lower prices, and higher welfare. This occurs because those regions are better export bases and thus attract firms.
- Multilateral trade liberalization reinforces competition in all regions. This generates richer product variety, higher productivity, lower prices, and higher welfare everywhere. Preferential trade liberalization increases productivity of insider countries while decreasing the average productivity of outsider ones. This translates in parallel changes in product variety, industrial activity, and welfare. The reason is that the liberalizing countries become better export bases: they gain better access to each other's market while maintaining the same ease of access to the third country's market. Average costs, prices, and markups move accordingly decreasing in the insiders and rising in the outsider countries.

Finally, Section .6 concludes by stressing the key message of the paper: trade liberalization induces a reallocation of resources from less to more productive firms, from smaller to larger countries, from high-cost to low-cost countries, and from outsiders to insiders in preferential trade agreements. This delivers long-run efficiency gains to liberalizing countries through selection among

heterogeneous firms, which eventually leads to higher average productivity, lower average prices, larger average firm size, higher profits, richer product variety and lower markups. At the same time, it generates tensions between prospective short-run winners (e.g. more efficient firms, larger and more developed countries, larger and more developed regions within countries, insiders in preferential trade agreements) and prospective short-run losers (e.g. less efficient firms, smaller and less developed countries, smaller and less developed regions within countries, outsiders in preferential trade agreements). Financial liberalization and labour market reforms can reduce the asymmetric distribution of gains and losses by speeding up the transition.

.2. Disparities in MERCOSUR

After almost fifteen years of MERCOSUR, one of the highest obstacles to deeper integration is increasingly represented by asymmetries between member countries. The issue has been brought to the forefront especially by smaller members due to their growing discontent about the distribution of MERCOSUR-related gains and losses. There is some substance to such dissatisfaction as asymmetries are real along many dimensions, such as economic size, individual wealth, specialization in production and patterns of trade. Moreover, there is little evidence that asymmetries have been narrowing since MERCOSUR started.

.2.1. Size and Wealth

As discussed by Moreira (2003), disparities among MERCOSUR countries are large in terms of both economic size and individual wealth. Along the first dimension, disparities are measured by comparing GDPs. Figure 1 shows that the difference between Brazil and the other partners is indeed huge. Brazil is more than thrice as big as Argentina and more than thirty times as big as Uruguay or Paraguay.

Individual wealth is measured, instead, by GDP per capita. In this respect, Figure 2 shows that countries are more commensurable. On average Uruguayans are the richest, being slightly richer than the Argentinians, one and a half times as rich as the Brazilians, and more than twice as rich than the Paraguayans. The Brazilian national figure itself hides deep internal differences between a richer South-East and poorer North-East.

.2.2. Production and Trade

Economic size disparities have an impact also on production and trade patterns. Sanguinetti, Traistaru and Volpe Martincus (2004) as well as Volpe Martincus (2004) show that Brazil is less specialized than other partner countries in terms of both production and trade patterns. Its industrial structure covers more evenly a richer set of industries, which allows for a more balanced

distribution of trade shares across sectors. Moreover, Brazil's trade pattern does not vary much between MERCOSUR and non-MERCOSUR countries. Other members are much more specialized and the trade shares of different industries are very different across destinations and reveal an increasing regional bias.

In terms of output composition, through the years MERCOSUR countries as a whole have changed by moving away from labour intensive sectors, such as textiles, wearing apparel and footwear, towards food products, chemicals, petroleum refineries, and transport equipment. Within MERCOSUR the reorganization of production has followed the comparative advantages of countries driven by the relative abundance of factors and natural resources. For example, Argentina and Uruguay have seen increased specialization in food and leather products; Brazil in wood and paper products; Argentina also in petroleum refineries; Brazil and Uruguay also in labour intensive footwear, textiles and wearing apparel. When countries specialize in sectors with no obvious ex-ante comparative advantage, state intervention is often the explanation.

As to trade with the rest of the world, the overall tendency is towards a decrease in specialization due the fact that sectors characterized by extreme comparative advantage are losing ground to sectors of milder comparative advantage. This is due to the rise of intra-industry trade.

.2.3. Intra-Industry Dynamics

Quite surprising is the fact that, apart from increased regional bias in trade, MERCOSUR per se seems to have had little impact on the aggregate economic behaviour of member countries. The reason is probably that, parallel to trade liberalization, other important policy reforms have been implemented in both labour and financial markets.

The separate impact of trade liberalization, however, can be detected when one looks at the micro level. Though studies based on firm-level data in MERCOSUR countries are still in limited availability, the few existing ones are starting to show some clear tendencies.

For instance, in the case of Argentina, Sanchez and Butler (2003) show that reforms have led to reallocations of resources that have taken place much more within than between sectors. That is, most of the observed structural change has involved the contraction of less productive establishments and the expansion of more productive ones within the very same industries. Moreover, among all policy reforms, trade liberalization is associated with the largest effect on reallocation and productivity followed by financial and labour reforms.

Casacuberta, Fachola and Gandelman (2004) study Uruguayan firms. They argue that, in principle, within an industry, trade liberalization may foster productivity through three main channels: improved access to foreign intermediate inputs and capital goods; tougher import competition pushing towards innovation; exit of the least productive firms. The dominance of this last channel is

stressed by the fact that liberalization has increased job creation, job destruction and capital destruction while it has lowered capital creation. Indeed, there is also some direct evidence that exiting firms tend to have lower productivity.

These examples nicely match a more general conclusion from recent studies made on several countries outside MERCOSUR (see, e.g., Tybout, 2002): the investigation of the microeconomics of competition and innovation within specific sectors is the research strategy that currently offers the highest return in terms of understanding both the origin and the destiny of asymmetries among countries involved in common integration processes. This calls for a theoretical approach that should be firm-based rather than sector-based.

To summarize:

Preliminary Observation – The separate impact of MERCOSUR on the asymmetries among member countries is hard to detect at the macro-level. Accordingly, the study of the microeconomics of competition and innovation within specific sectors is necessary to understand the origin and the destiny of asymmetries among countries involved in common integration processes.

.3. *Agglomeration Forces as the Source of Disparity*

When observing geographical asymmetries in economic development, the first obvious explanation that comes to one's mind is that regions differ in terms of their relative abundance of natural resources, their proximity to natural means of communication, and their climatic conditions. All these characteristics define the exogenous attributes of a region, hence what Cronon (1991) calls 'first nature', and they play centre stage in traditional trade theories of comparative advantage along the lines drawn by Ricardo, Heckscher and Ohlin. In particular, those theories argue that:

Insight 1 – International cost differences foster the concentration of industries in countries where the corresponding costs are lower.

For a specific sector, these are regions that: (i) use relatively advanced technologies in the sector; (ii) are relatively abundant in the factor in which the sector is relatively intensive; (iii) offer better local infrastructures for transporting intermediate goods.

However, dramatic differences in economic development can be observed even between areas that are not very different in terms of those exogenous attributes. This suggests that the observed

regional unbalances must be driven by some other forces ('second nature') that are inherent to the functioning of economic interactions and that, in principle, are able to generate uneven development even across ex-ante identical places.

As discussed by Fujita and Thisse (2002), second nature explanations have a long history in economics, geography, and regional science. However, during the last ten years, the debate within mainstream economics has been dominated by NEG. With respect to alternative approaches, the defining feature of NEG is its focus on market rather than non-market interactions. This is pursued within a 'general equilibrium' framework stressing the endogenous determination of good and factor prices and the importance of economy-wide budget constraints. In the words of Fujita and Krugman (2004): "you want a general-equilibrium story, in which it is clear where the money comes from and where it goes". The aim of the present section is to illustrate the theoretical foundations of NEG.

.3.1. A Firm-Based Approach

Being based on market interactions, NEG places the location decision of the firm at the heart of its approach. Such decision is not trivial when two things are true. First, goods and factors can be transported across space only at some cost. Second, the fragmentation of the production process reduces its efficiency, which happens when returns to scale are increasing at the plant level. Without transport costs space would be immaterial. Without plant-level scale economies, when faced with dispersed customers and suppliers, firms would use the geographical fragmentation of production to circumvent transport costs by patronizing scattered demand and intermediate supply through many small local plants ('backyard capitalism'). Thus, both transport costs and scale economies are necessary for a location problem to arise: costly transportation gives physical substance to the concept of geography and increasing returns generate an economic trade-off between 'proximity' to customers and 'concentration' of production in as few plants as possible. Scotchmer and Thisse (1992) call this the 'folk theorem of spatial economics'.

The centrality of scale economies has important implications in terms of market structure. Indeed, as pointed out by Starrett (1978), since plant-level returns to scale are necessarily associated with market power, imperfect competition is inherent to the location problem. Indeed, the tension between proximity and concentration generates a location problem for any firm no matter whether involved in any interaction with other firms. The problem, however, becomes more complicated once firms face competitors. The reason is that geographical positioning can be used by firms to relax competitive pressures and enhance market power. Using Chamberlin's (1933) terminology, this is the case both when competitors are a 'small group' (oligopoly) and when they are a 'large group' offering differentiated products (monopolistic competition). In both cases, location turns out

to be a crucial decision variable for profit maximization as it allows firms to increase their market power by careful positioning.

On the one hand, firms are attracted towards markets with large local demand. The reason is that, by locating close to customers, firms are able to save on trade costs. Hence:

Insight 2 – International differences in local market size foster the agglomeration of industries in larger countries.

This is sometimes called the ‘home market effect’ (Krugman, 1980; Helpman and Krugman, 1985), whereby firms tend to solve the trade-off between proximity and concentration by serving the smaller market from the larger one.

On the other hand, since markets with large local demand attract firms, they are likely to repel firms. The reason is that market power is hampered by the presence of many competitors, so firms can increase their profits by fleeing markets with a high density of firms in their same sectors. The more so, the more similar competing products are. Accordingly:

Insight 3 – Product market competition promotes the geographical dispersion of industries.

Through imperfect competition, plant-level scale economies have also crucial implications in terms of welfare. The reason is that, when firms have market power, the prices, on which consumers and firms base their consumption, production and location decisions, do not fully reflect the corresponding social values. This means that market interactions generate ‘side effects’ for which no quid-pro-quo is paid. Being associated with market transactions, those ‘side effects’ are called ‘pecuniary externalities’ (Scitovsky, 1954).

Three possible scenarios are especially relevant. In the first, when a firm relocates, it decreases competition in the place of origin and increases competition in the place of destination. Accordingly, the profits of competitors rise in the former and fall in the latter. A pecuniary externality materializes in both places in so far as the relocating firm disregards those effects. In particular, the relocating firm imposes a positive externality on its competitors in the place of origin and a negative externality on its competitors in the place of destination. By pushing down profits in places crowded by firms, competition therefore acts as a dispersion force (Ottaviano, Tabuchi and Thisse, 2002). The second scenario considers the effect of firm relocation when matched by labour migration. In this case, as the firm moves, it reduces demand in the place of origin while increasing it in the place of destination. In so doing, as profits rise with demand, the firm imposes a negative externality on

competitors in the former place and a positive one on competitors in the latter. By raising profits in places crowded by firms, market size therefore acts as an agglomeration force (Krugman, 1991). In the third scenario, firms are linked by input-output linkages: what is output for a firm is input for the others and vice versa. Here, when a firm relocates, it depresses both final demand and intermediate supply in the country of origin, whereas it reinforces them in the country of destination. Accordingly, other firms' profits suffer in the former country, where the firm imposes a negative externality, and thrive in the latter, where it imposes a positive externality. By raising profits in places crowded by firms, input-output linkages therefore act as an agglomeration force (Krugman and Venables, 1995; Venables, 1996).

Differently from NEG, other approaches stress the role of 'technological' rather than 'pecuniary' externalities (see, Marshall, 1890, Henderson, 1978, as well as Ciccone and Hall, 1996, for recent reassessments). Technological externalities differ from pecuniary ones in that they materialize through sheer physical proximity independently from any market transaction (Scitovsky, 1954). As they arise from non-market interactions, also for them no quid-pro-quo is paid. In the textbook case, the productivity of a firm is influenced by the presence of other firms nearby even without any market relation with them. As pecuniary externalities, also technological ones can be either negative or positive. On the one hand, nearby firms may reduce a firm's productivity through the pollution they generate or through the congestion they cause in the use of local public goods and infrastructures. On the other hand, nearby firms may increase a firm's productivity through informal knowledge transmission ('spillover') generated as a by-product of their contacts with the surrounding environment.

To summarize, the geographical distribution of demand and the position of other firms determines the relative attractiveness to a firm of alternative locations through market or non-market interactions. This creates a feedback mechanism among firms' location decisions through which firms' interactions ('second nature') may alter the economic landscape implied by natural resources, natural means of communication, and climatic conditions ('first nature'). In addition, since 'second nature' is driven by localized externalities, in a free market the location of firms is generally inefficient and appropriate public intervention is generally needed.

.3.2. Micro-Founded Agglomeration

The results that the location decisions of firms are intertwined and this may generate disparities and inefficient location follow from both pecuniary and technological externalities. However, the former externalities do have a logical advantage with respect to the latter, which lies in the possibility of relating their emergence to a set of well-defined microeconomic parameters. This has proven to be quite difficult in models based on technological externalities as these still remain

mostly ‘black boxes’ (see, e.g., Ottaviano and Thisse, 2001, as well as Duranton and Puga, 2004, for recent assessments).

To explain this point, let us expand one of the previous examples that involved input-output linkages. In particular, following Venables (1996), consider a production chain consisting in three vertically linked activities: intermediate production, final production, and consumption. For simplicity, assume that final production uses only intermediate inputs, intermediate production employs only labour, workers are the only source of final demand and they are geographically immobile. If, for any reason, a new firm starts producing intermediates, it will increase labour demand and intermediate supply. Due to excess demand and supply respectively, wages will go up while intermediate prices will fall. This is bad news for the other intermediate producers due to competitor proximity. However, it is good news for final suppliers, who experience falling production costs and higher demand by richer workers. As new final producers are lured into the market, the expansion of final production will feed back into stronger intermediate demand so that also intermediate suppliers will benefit thanks to customer proximity. Clearly, when the latter effect dominates the former, both final and intermediate firms will end up being agglomerated in the same place. Accordingly, circular causation among firms’ location decisions can generate persistent differences even among initially identical places (‘second nature’).

The crucial contribution of NEG is that such simple arguments are translated into general equilibrium models with solid microeconomic foundations. This allows the evolution of the spatial landscape to be related to observable microeconomic parameters. NEG predicts agglomeration to be more likely to take place in sectors where market power is strong. This is the case when there are intense plant-level scale economies and pronounced product differentiation. The reason is that stronger market power weakens the dispersion effect of competition. To put it differently, market power gives strength to ‘second nature’ against ‘first nature’, which detaches the emerging economic landscape from the physical attributes of its underlying geography. Thus, a priori there is great flexibility on where particular activities locate. Nevertheless, once the agglomeration process has started, spatial differences take shape and, as localized pecuniary materialize, economic geography becomes quite rigid. This is what Fujita and Thisse (1996) call ‘putty clay’ geography.

To summarize:

Insight 4 – Increasing returns and product differentiation enhance the attractiveness of larger markets to both firms and workers. As firms and workers move, those markets become even larger, which may generate cumulative agglomeration processes.

For example, in the presence of local spillovers, country-specific cost differences may arise endogenously from the spatial concentration of firms (see, e.g., Martin and Ottaviano, 2001). Analogously, in the presence of migration (Krugman, 1991; Ottaviano, Tabuchi and Thisse, 2002) and capital accumulation (Baldwin, 1999; Baldwin, Martin and Ottaviano, 2001), relative factor endowments and market sizes may be endogenously determined by firm clustering. Lastly, in the presence of input-output linkages, the overall transport bill may be reduced by the spatial concentration of customers and suppliers (Krugman and Venables, 1995; Venables, 1996).

.3.3. Attraction and Accessibility

Together with the role of plant-level scale economies and product differentiation in generating localized externalities, arguably the most celebrated insight of NEG is the impact of transportation improvements and trade liberalization on the economic landscape. The reason is that, differently from other approaches, micro-foundation allows NEG to reach a deeper understanding of how economic geography changes as trade impediments are gradually eliminated. In particular, NEG argues that the level of trade barriers affects the balance between the agglomeration push of market size and the dispersion pull of competition.

With extremely high trade barriers, competition leads to dispersion as markets can be reached only through local production. Trade liberalization, however, weakens the dispersion force due to competition more than it weakens the agglomeration force due to international size differences. This promotes agglomeration as trade barriers are reduced from the initially high levels. At the same time, as trade barriers keep on falling, both forces vanish. What is eventually left is the agglomeration force due to international cost differences, so dispersion may reappear for low level of trade costs.

Insight 5 – Trade liberalization has a non-linear effect on the spatial concentration of economic activities by promoting agglomeration at early stages and dispersion afterwards.

This behaviour is sometimes called the ‘bell-shaped curve of spatial development’ (Ottaviano and Thisse, 2004).

In a realistic set-up with many countries, the basic concept underlying the analysis is the so-called ‘market potential’ (Harris, 1954). This has both nominal and real definitions (Head and Mayer, 2004). Whereas the ‘nominal market potential’ (henceforth, NMP) is a measure of customer proximity, the ‘real market potential’ (henceforth, RMP) is a combined measure of customer and competitor proximity. Formally, consider a group of locations. The nominal market potential of a certain location H is the weighted average expenditures across all locations that plants can tap if

located in H. Differently, the real market potential of H is the weighted average real expenditures (‘purchasing power’) across all locations that plants can tap if located in H. In both cases, the weight of each location is a decreasing function of its distance from H. The underlying idea is that NMP is a good proxy of the value of sales that plants can expect to make on average if located in H. Differently, RMP is a good proxy of the profits than an average firm can make if located in H. In the long run, since firms can freely pick plant locations, profits should reach the same normal level everywhere. Therefore, in the long run RMP differences should eventually vanish as NMP differentials are capitalized in local price differences. Accordingly, short-run RMP differences should predict the future evolution of the economic landscape as firms are attracted towards areas temporarily boasting higher RMP.

As pointed out by Behrens et al (2004), the concepts of NMP and RMP are closely related to spatial interaction theory. The NMP of a certain area captures both the size of its local market (‘attraction’) and its connection to other markets (‘accessibility’). In addition, the RMP captures the intensity of competition faced by firms located in that area (‘repulsion’). Attraction, accessibility and, to a lesser extent, repulsion are also the main ingredients of gravitational models of international trade (Head and Mayer, 2004).

To summarize, according to NEG:

Insight 6 – The evolution of the economic landscape is driven by relative changes in the market potentials of alternative locations.

As the market potential of a location is determined by its attraction and accessibility, any changes in the overall distribution of country sizes and/or in the exchange network is bound to affect the spatial allocation of economic activities. As discussed in Section .5, this is crucial for regional trade agreements like MERCOSUR.

.4. Trade Liberalization, Agglomeration and Firm Selection

According to traditional theories, trade liberalization allows countries to specialize in sectors in which they enjoy a comparative advantage, because of better technology or suitable factor endowments. This is due to the possibility of buying in international markets products that are no longer supplied by domestic firms. Specialization takes place through the reallocation of productive factors from sectors that are relatively high-cost to sectors that are relatively low-cost in terms of international standards.

As discussed in the Section .1, this sort of inter-sector reallocation is taking place in MERCOSUR. However, any relevant separate impact of regional trade liberalization itself is hard to spot. On the contrary, trade integration seems to be the main driving force behind very visible intra-sector reallocations. This matches the existing international literature on episodes of rapid liberalization as surveyed by Tybout (2002). Such literature concludes that within sectors trade liberalization forces less productive firms out of the market, thereby increasing average productivity. This is called the ‘selection effect’ of international trade and is accompanied by higher average firm size (‘scale effect’) as well as lower average price and mark-up (‘pro-competitive effect’)

.4.1. Market Potential Matters

To understand the basic mechanism behind the selection effect, let us consider what happens to firms’ profits after trade liberalization (see, e.g., the models by Melitz, 2003; Bernard, Redding and Schott, 2004; Ottaviano and Melitz, 2005).

Let us focus on a sector that is active in two autarchic countries that are identical in terms of preferences, technologies and factor endowments. Firms are differentiated horizontally in terms of the products they sell and vertically in terms of their productivities (or, equivalently, in terms of the qualities of their products). Firms have market power and this is stronger for more productive firms, which are able to sell at lower prices, quote lower mark-ups but still make higher profits due to larger size. For given productivity, market power is also positively related to the total number of competitors, so all mark-ups fall as the number of competing firms increases. Finally, there are increasing returns to scale at the firm level due to fixed cost and entry as well as exit are free. This implies that the least productive firms are just able to break even as their fixed costs are just covered by the operating profits associated with positive mark-ups.

Consider now what happens when trade is liberalized. In so doing, we distinguish between the ‘short run’, when firms are not able to enter or exit the market, and the ‘long run’, when entry and exit take place. How long is the long run depends on financial and labour market institutions whose rigidities may hamper capital as well as job destruction and creation.

As the two countries go from autarchy to free trade, the two national markets merge into a common international market. Since countries are identical, initially in the new market there are twice as many firms as in each autarchic market, so mark-ups have to fall. In the short run, firms cannot enter or exit, so operating profits fall for all firms no matter how productive they are. However, this is a more serious problem for the least productive firms. Since they were just breaking even in autarchy, they are now making losses. This implies that in the long run their position in the market will become unsustainable and they will have to exit. Their lost market shares will be captured by more productive firms, which will more than compensate the fall in their mark-ups by larger scale.

In the end, average productivity and firm size will be higher, whereas average mark-up and thus price will be lower. The reason is the ‘survival of the fittest’ in a tougher competitive environment and selection will be harsher when market power is not shielded by product differentiation and increasing returns to scale.

These results can be also interpreted as implying that a larger market or a market with lower behind-the-border restrictions is more selective and therefore supports a more productive group of firms. Hence:

Insight 7 – Larger integrated markets are characterized by firms that on average are more productive, operate at a larger scale and quote lower prices. The more so the less differentiated products are and the weaker the returns to scale.

This result can be extended by analogy when liberalization falls short of free trade so that some additional trade barriers still persist, for instance due to additional costs of transport and distribution for foreign sales. As already discussed, when some export costs persist, the intensity of local competition does not depend only on the ‘attraction’ of a country’s domestic market but also on its ‘accessibility’ to and from other countries. Insight 7 then leads to:

Insight 8 – Countries with higher market potential host firms that on average are more productive, operate at a larger scale and quote lower prices. This happens especially in sectors with little product differentiation and weak returns to scale.

.4.2. Trade Integration and Financial Liberalization

When liberalization does not dismantle all trade barriers, the selection effect gets richer. When reform leads the economy from autarchy to free trade, all surviving firms sell to all customers, not only domestic but also foreign ones. Residual trade barriers cause, instead, a second round of selection within the group of surviving firms, called ‘selection into export status’. The reason is that two effects are at work. On the one side, lower trade barriers make domestic markets easier to accede by foreign firms. This creates import competition, which cuts into firms profit. On the other hand, lower trade barriers also make it easier for domestic firms to reach foreign markets and improved export access increases their profits. The former effect is bad news for all firms no matter how productive they are. Differently, the second channel is good news for some firms only: those that are productive enough to cover the additional export costs with their fatter operating profits. Accordingly, the firms that were active in autarchy, end up being partitioned by trade liberalization

into three groups. The ones with low productivity leave the market. The very productive ones sell in both their domestic and foreign markets. Those with medium productivity sell only in their domestic market. Hence:

Insight 9 – Trade liberalization causes the exit of the least productive firms and the selection of the most productive survivors into export status. Other surviving firms are confined to their domestic markets.

As a result, trade liberalization increases average productivity as well as firm size and decreases average prices. The more so the less differentiated products are and the weaker the returns to scale. As already mentioned, the speed of all these effects depends on the functioning of financial and labour markets as these mediate the reallocation of capital and labour from hailing to thriving firms. They also foster innovation and the creation of new firms. In this respect, for emerging countries that face a shortage of internal capitals, it may be crucial to match trade integration with financial liberalization: international capital flows and FDI can smoothen the process of adjustment through selection. Thus:

Insight 10 – Financial liberalization and FDI speed up the reallocation of resources from less productive to more productive firms triggered by trade liberalization.

The credibility of trade liberalization and the associated reduction of uncertainty about possible reversals of the process would have a similar smoothing effect by making workers and firms more willing to incur the costs of adjustment.

.4.3. Comparative Advantage

Once we move away from autarchy, firm selection is affected not only by market potential but also by comparative advantage. Intuitively, for given size, a more technologically advanced country represents a tougher competitive environment. Due to stronger selection effects it will host firms that on average have higher productivity, reach larger scale and quote lower prices. In this sense, selection magnifies comparative advantage. Thus:

Insight 11 – In sectors where they enjoy a comparative advantage, countries host firms that on average are more productive, operate at a larger scale and quote lower prices.

As an example, Figures 3, 4, 5, 6, and 7 show the impact of trade liberalization between a larger country H and a smaller country F. The focus is on a sector in which country H has a comparative advantage, so H offers both larger market size and lower production costs. For each country the figures respectively depict the maximum feasible production cost (i.e. the ‘cutoff’ cost of the least productive firms that just break even), the numbers of sellers, entrants, and producers, as well as the levels of welfare in the two countries as functions of trade barriers. Solid lines refer to country H, dashed lines to country F, and, when present, dotted lines to the whole economy. The figures point out that, even though the small backward region loses its industrial base as trade gets freer, welfare levels nonetheless converge since the location of producers becomes progressively immaterial.

.5. *Integration and Disparities: a Numerical Example*

To add some concreteness to the foregoing insights, it is interesting to present the results of a numerical investigation of a stylized economy embedding all the effects highlighted in previous sections. The model is adapted from Melitz and Ottaviano (2005).

The main idea discussed above is that three features shape the attractiveness of a country to firms: locally available technologies and inputs (‘cost-saving attraction’), the size of the local market for final products (‘market-seeking attraction’), the access to and from foreign markets (‘accessibility’). In turn, the interactions among those three features determine the individual characteristics of local firms and these determine the wealth of the country. Using the terminology introduced in Section .3 technologies and inputs determine the ‘first nature’ of the country whereas the size of the local market and the access to foreign markets shape its ‘second nature’ in terms of the technologies that are actually adopted. This happens through a selection process that sets a lower bound on the productivity (i.e. a higher bound on the marginal costs) of firms that are able to operate locally. Most naturally, trade agreements affect accessibility and, therefore, the selection process.

.5.1. *Attraction and accessibility*

Consider six countries and nine regions. Regions are called LA, SA, CO, BA, AS, MO, PA, SP, BR. CO and BA belong to the same country. PA, SP and BR to another country. All other regions coincide with their own countries.

Labour is the only factor of production and there are two industries. In one industry workers are employed by firms with no market power that produce a freely traded homogeneous good under constant returns to scale. In the other industry labour is employed by firms with market power, each producing a variety of a horizontally differentiated good under increasing returns to scale. Regions differ in terms of size, accessibility and comparative advantage as illustrated in Figures 8 and 9. The former reports the hierarchy of regions in terms of size. The latter figure shows the hierarchy in

terms of a joint measure of accessibility and comparative advantage. For region H such measure is constructed as the weighted average of the maximum possible production costs across regions with weights determined by the relative position of regions within the trade network. Specifically, the production cost of each region enters the average with a weight that is inversely related to its trade costs with H. All the rest given, the higher this measure is with respect to other regions, the more competitive firms in region H are likely to be (see Melitz and Ottaviano, 2005, Appendix B, for details).

In general, trade costs include all distance-related costs such as transport costs, administrative barriers and currency conversion costs. Figure 9, however, abstracts from international trade barriers and currency conversion costs in order to focus on the inertial part of trade costs that cannot be instantaneously changed by *fiat* and is thus less responsive to policy shifts. It shows that, whereas regions BA, PA and SP stand out in terms of size (see Figure 8), their lead is challenged by BR and CO when it comes to accessibility and comparative advantage.

.5.2. Integration scenarios

Four scenarios are investigated: no trade integration ('NO'); preferential trade integration excluding LA and SA ('MER'); enlarged integration ('ENL') when trade integration is extended also to LA and SA; monetary union when MER countries also share a common currency ('MU'). These scenarios imply different trade costs, which are constructed by using the values reported by Anderson and van Wincoop (2004). These authors provide a 'consensus estimate' of the average trade cost for developed countries (and a lower bound for the average trade cost for developing countries) and their decomposition in local distribution costs, transportation costs, border costs. These are further decomposed into components stemming from policy barriers, different currencies, different languages, information barriers, and security barriers.

The outcomes of the different trade reforms are shown in Figures 10-14. Figure 10 depicts the cutoff costs (i.e. an inverse measure of average productivity) under the various scenarios. Without trade integration, average productivity is higher in places blessed by large local markets, good accessibility and comparative advantage. Preferential trade liberalization improves the average productivity in member countries while reducing it in excluded ones. A common currency gives an additional boost to members' productivity while having negligible effects on excluded countries. Enlarged integration makes average productivity grow in the new member countries while having no impact on old ones. Mirror results are reported in Figure 11 for product variety.

Figures 12 and 13 show that the somewhat limited asymmetries in terms of products available in each local market hide extreme asymmetries in industry dynamics. Large, low-cost and accessible markets see most of the action in terms of both entry and production. Trade liberalization boosts

industrial activity only in PA and SP thanks to their large local markets. Enlargement promotes industry in the old members while putting pressure on the industrial bases of the new members. The adoption of a common currency has a positive impact on entry in the larger regions but little impact on production patterns.

Figure 14 completes the overall picture by showing that, notwithstanding the geographical polarization of industry, no region loses and most regions gain from preferential trade liberalization, from enlargement, and from the adoption of a common currency. This stems from the fact that, in an integrated economy, the actual location of production loses importance with respect to cheap access to a wide variety of goods wherever they are produced.

Before concluding, two issues are worth commenting upon. First, if one cared about the actual location of producers, a key insight is that the initial disadvantage in size could be compensated by advantages in comparative costs and accessibility. This has important implications for small regions. Using the current numerical example, one can calculate by how much small regions need to improve their initial cost-saving attraction and accessibility in order to compensate for the disadvantage of being small. Consider, for instance, AS and MO, which have small size (Figure 8) but fairly good cost-saving attraction and accessibility (Figure 9). Starting with no trade integration ('NO'), liberalization leads to losses in the number of producers of 30 per cent in AS and 20 per cent in MO. These losses could be avoided if the joint measure of accessibility and comparative advantage increased by 5 and 7.5 per cent respectively relative to other regions.

Second, the figures show only the long run effects, that is, what happens after firms have had enough time to enter or exit the market (see Section .4.1). As reported by Melitz and Ottaviano (2005), notable deviations would take place in the short run only if liberalization were not simultaneously undertaken by all countries. In particular, countries that liberalize unilaterally would be hurt in the long but not in the small run when entry and exit are restricted. The same would happen for outsiders in preferential trade agreements. Therefore, in the presence of initially asymmetric trade liberalization the speed of completion of the integration process is crucial: slow processes sustain asymmetric gains and losses that fast processes do not.

.6. Conclusion

What are the obstacles ahead of MERCOSUR? Can the distribution of the associated costs and benefits create resistance from national interests to move towards deeper integration?

To answer these question, the present paper has proposed a theoretical framework to assess the economic impact and the welfare implications of the resulting reallocation of resources across firms

and countries. The proposed framework merges the key insights of ‘new economic geography’ with more recent developments on the selection effects of trade liberalization.

The main result is that trade liberalization induces a reallocation of resources from less to more productive firms (‘selection’), from smaller to larger regions (‘market-seeking attraction’), from high to low cost regions (‘cost-saving attraction’) and from outsiders to insiders in preferential trade agreements (‘accessibility’). This delivers long-run efficiency gains to liberalizing countries as tougher selection leads to higher average productivity, lower average prices, larger average firm size, higher profits, richer product variety, and lower markups. At the same time, it generates tensions between short-run winners and short-run losers by putting pressure on small remote backward countries and, within countries, on small remote backward regions as well as on low-productivity firms and workers. Under this respect, financial liberalization, labour market reforms as well as an improvement in the credibility of further trade liberalization can relax those tensions by reducing the costs of adjustment.

References

- Anderson, J.E. and E. van Wincoop (2004) Trade costs, *Journal of Economic Literature* 42, 691-751.
- Baldwin, R. (1999) Agglomeration and endogenous capital, *European Economic Review* 43, 253-280.
- Baldwin, R., P. Martin and G. Ottaviano (2001) Global income divergence, trade and industrialization: The geography of growth take-off, *Journal of Economic Growth* 6, 5-37.
- Baldwin, R., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nicoud (2003) *Economic Geography and Public Policy* (Princeton: Princeton University Press).
- Behrens, K., A. Lamorgese and T. Tabuchi (2004) Testing the home market effect in a multi-country world: The theory, CEPR, Discussion Paper No. 4468.
- Bernard, A., S. Redding and P. Schott (2004) Comparative advantage and heterogeneous firms, NBER, Working Paper No. 10668.
- Casacuberta C., G. Fachola and N. Gandelman (2004) The impact of trade liberalization on employment, capital and productivity dynamics: Evidence from the Uruguayan manufacturing sector, Inter-American Development Bank, Latin American Research Network, Working Paper No. R-479.
- Chamberlin, E. (1933) *The Theory of Monopolistic Competition* (Cambridge (Mass.): Harvard University Press).

- Ciccone, A. and R. Hall (1996) Productivity and the density of economic activity, *American Economic Review* 87, 54-70.
- Cronon, W. (1991) *Nature's Megalopolis: Chicago and the Great West* (New York: Norton).
- Duranton, G. and D. Puga (2004) Micro-foundations of urban agglomeration economies. In V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, vol. 4 (Amsterdam: Elsevier).
- Fujita, M. and P. Krugman (2004) The new economic geography: Past, present and future, *Papers in Regional Science* 83, 139-164.
- Fujita, M. and J.-F. Thisse (1996) Economics of agglomeration, *Journal of the Japanese and International Economies* 10, 339-378.
- Fujita, M. and J.-F. Thisse (2002) *Economics of Agglomeration: Cities, Industrial Location and Regional Growth* (Cambridge: Cambridge University Press).
- Fujita, M., P. Krugman and A. Venables (1999) *The Spatial Economy: Cities, Regions and International Trade* (Cambridge (Mass.): MIT Press).
- Head, K. and T. Mayer (2004) The empirics of agglomeration and trade. In V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, vol. 4 (Amsterdam: Elsevier).
- Helpman, E. and P. Krugman (1985) *Market Structure and Foreign Trade* (Cambridge (Mass.): MIT Press).
- Henderson, V. (1978) *Economic Theory and the Cities* (London: Academic Press).
- Krugman, P. (1980) Scale economies, product differentiation and the pattern of trade, *American Economic Review* 70, 950-959.
- Krugman, P. (1991) Increasing returns and economic geography, *Journal of Political Economy* 99, 483-499.
- Krugman, P. and A. Venables (1995) Globalization and the inequality of nations, *Quarterly Journal of Economics* 110, 857-880.
- Marshall, A. (1890) *Principles of Economics* (Macmillan: London).
- Martin, P. and G. Ottaviano (2001) Growth and agglomeration, *International Economic Review* 42, 947-968.
- Melitz, M. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica* 71, 1695-1725.

- Melitz, M. and G. Ottaviano (2005) Market size, trade, and productivity, NBER, Working Paper No. 11393
- Moreira, M. (2003) A short note on MERCOSUR's asymmetries, Inter-American Development Bank, mimeo.
- Neary, P. (2001) Of hype and hyperbolas: Introducing the new economic geography, *Journal of Economic Literature* 39, 536-561.
- Ottaviano, G. (2003) Regional policy in the global economy: Insights from New Economic Geography, *Regional Studies* 37, 665-674.
- Ottaviano, G. and D. Puga (1998) Agglomeration in the global economy: A survey of the 'new economic geography', *World Economy* 21, 707-731.
- Ottaviano, G. and J.-F. Thisse (2001) On economic geography in economic theory: increasing returns and pecuniary externalities, *Journal of Economic Geography* 1, 153-179.
- Ottaviano, G. and J.-F. Thisse (2004) Agglomeration and economic geography. In J.V. Henderson and J.-F. Thisse (eds.) *Handbook of Regional and Urban Economics*, Vol.4 (Amsterdam: Elsevier).
- Ottaviano, G., T. Tabuchi and J.-F. Thisse (2002). Agglomeration and trade revisited, *International Economic Review* 43, 409-435.
- Ottaviano, G. and D. Pinelli (2004) The challenges of globalization for Finland and its regions: The new economic geography perspective, Prime Minister's Office Publications, 24/2004, Helsinki.
- Sanchez, G. and I. Butler (2003) Market institutions, labor market dynamics, growth and productivity: Argentina, IERAL, mimeo.
- Sanguinetti, P., I. Traistaru and C. Volpe Martincus (2004) The impact of South-South preferential trade agreements on industrial development: An empirical test, *Universitat Torquato di Tella*, mimeo.
- Scitovsky, T. (1954) Two concepts of external economies, *Journal of Political Economy* 62, 143-151.
- Scotchmer, S. and J.-F. Thisse (1992) Space and competition: A puzzle, *Annals of Regional Science* 26, 269-286.
- Starrett, D. (1978) Market allocations of location choice in a model with free mobility, *Journal of Economic Theory* 17, 21-37.

- Tybout, J. (2002) Plant and firm-level evidence on new trade theories. In J. Harrigan (ed.) Handbook of International Economics (London: Blackwell Publisher)
- Venables, A. (1996) Equilibrium locations of vertically linked industries, International Economic Review 37, 341-359.
- Volpe Martincus, C. (2004) Changing specialization patterns in MERCOSUR, ZEI Bonn, mimeo.

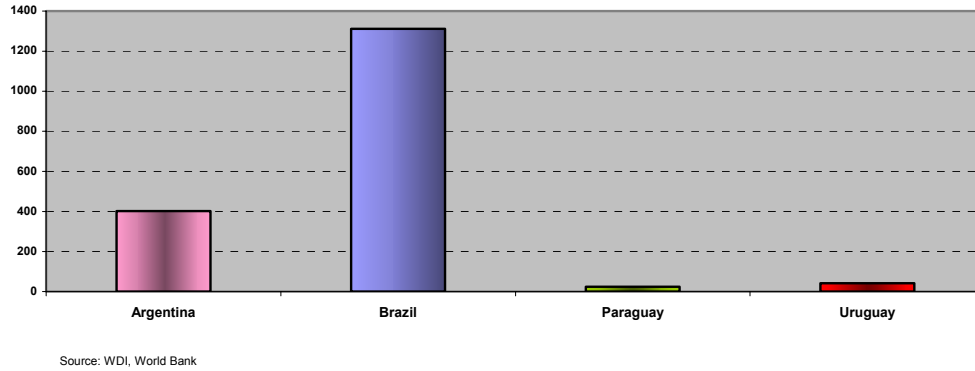


Figure 1 - GDP, 2002 PPP (current international \$)
Source: Moreira (2003)

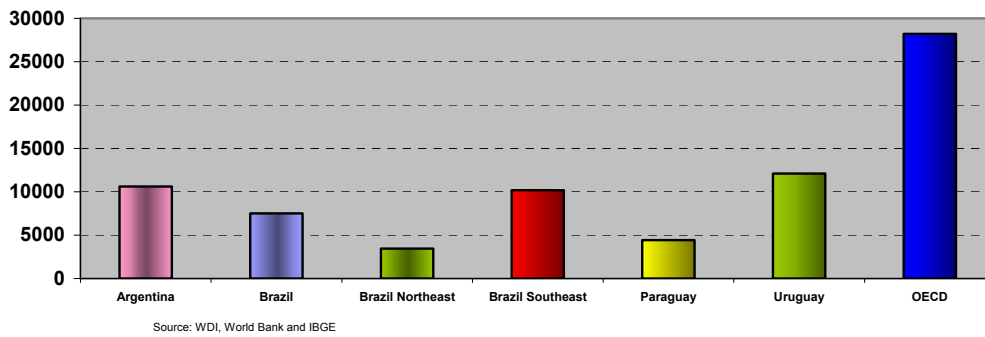


Figure 2 - GDP per capita, 2002 PPP (current international \$)
Source: Moreira 2003

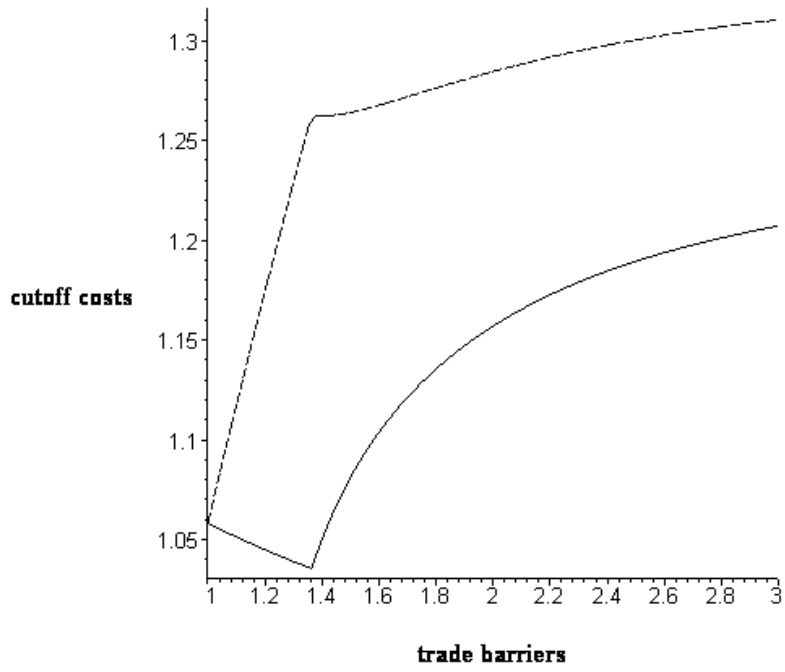


Figure 3 – Two countries: Selection

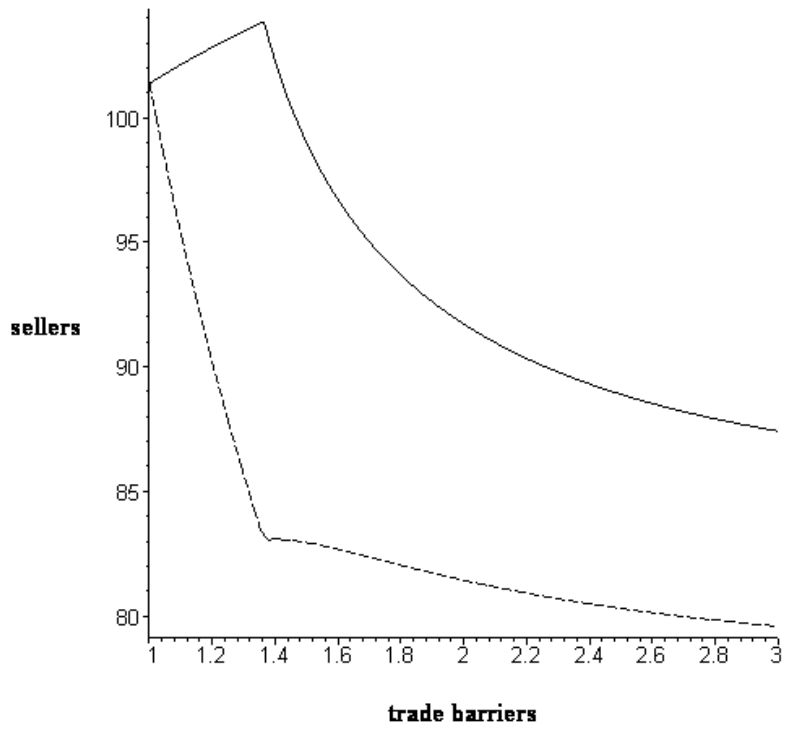


Figure 4 – Two countries: Variety

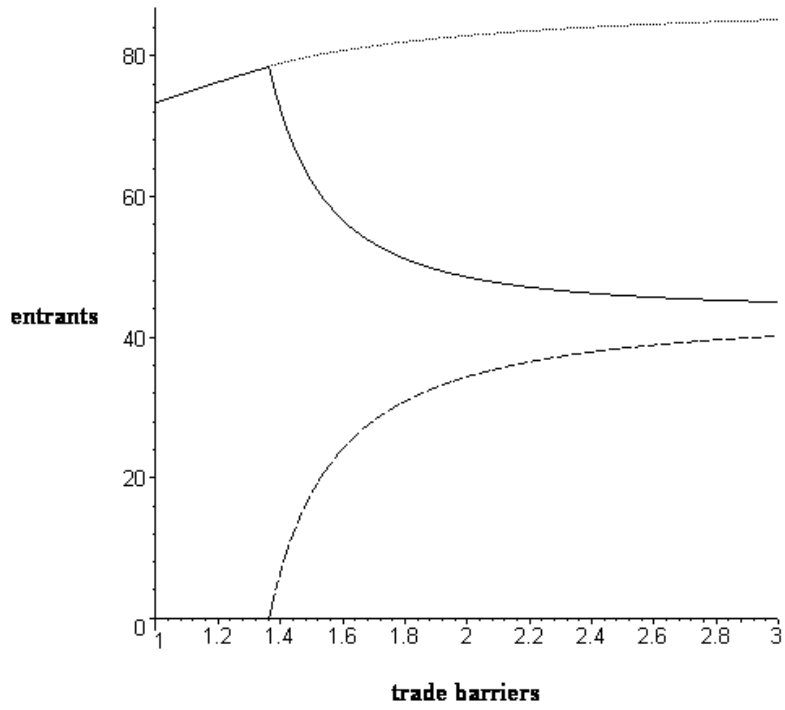


Figure 5 – Two countries: Entry

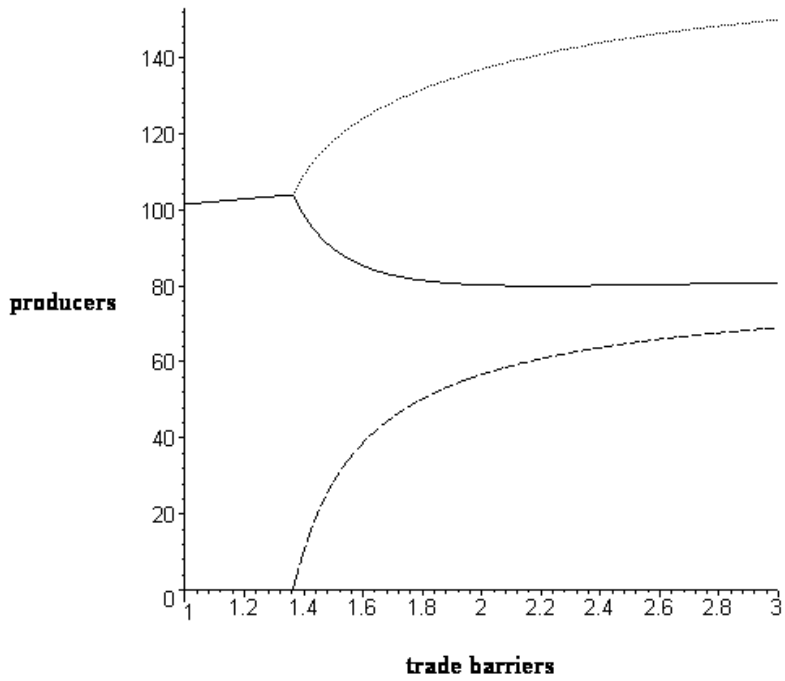


Figure 6 – Two countries: Production

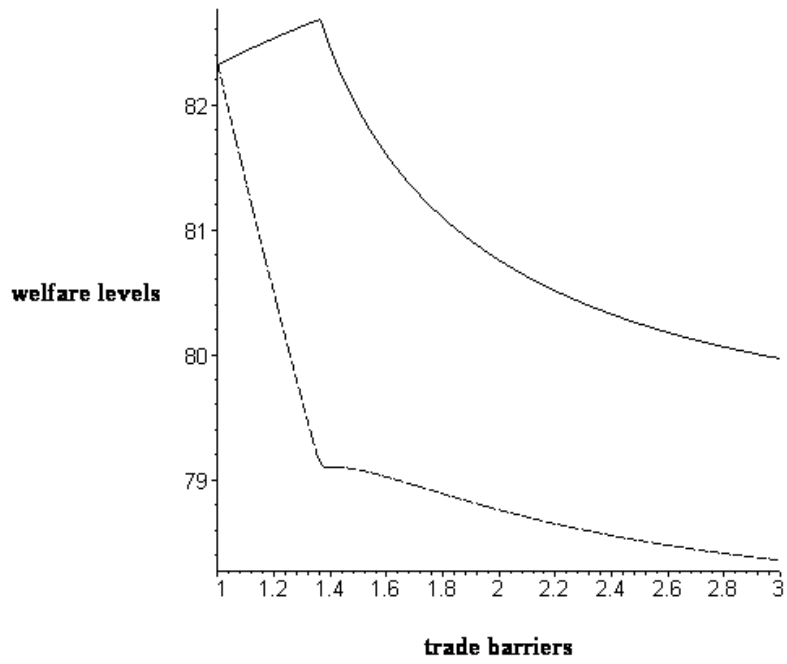


Figure 7 – Two countries: Welfare

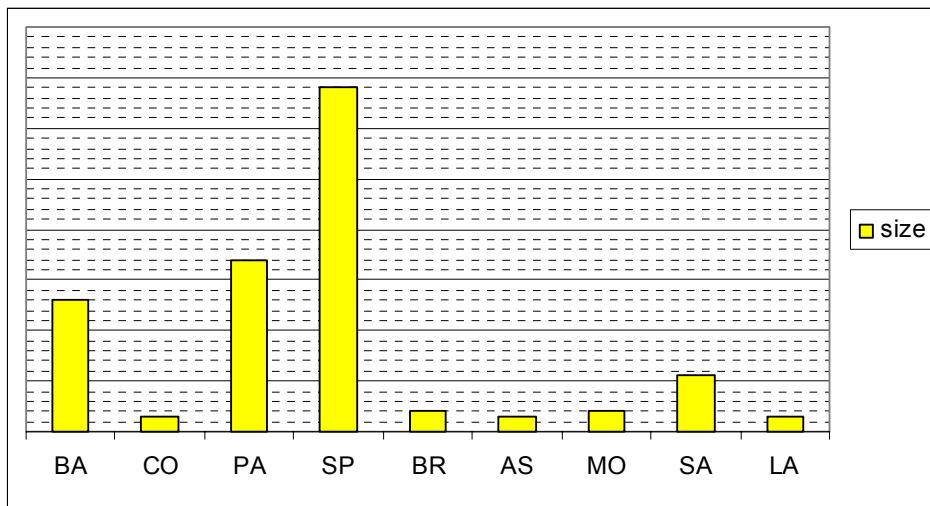


Figure 8 – Regional integration: Market-seeking attraction

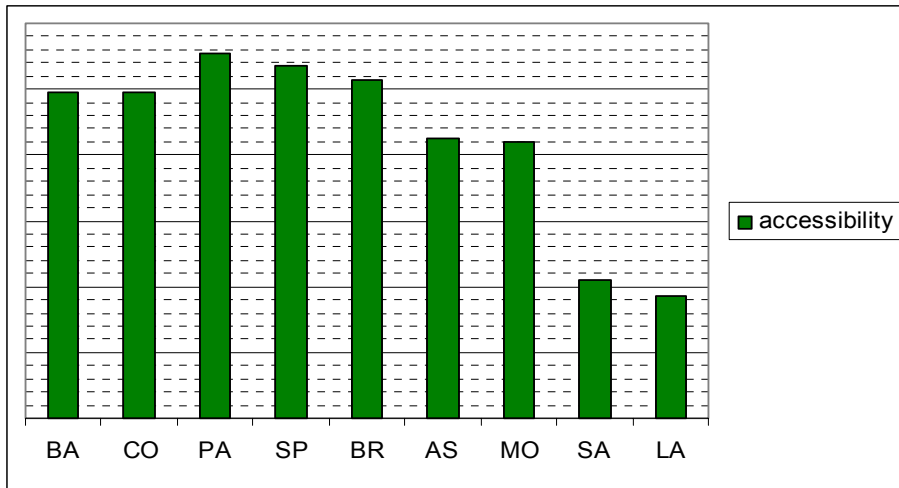
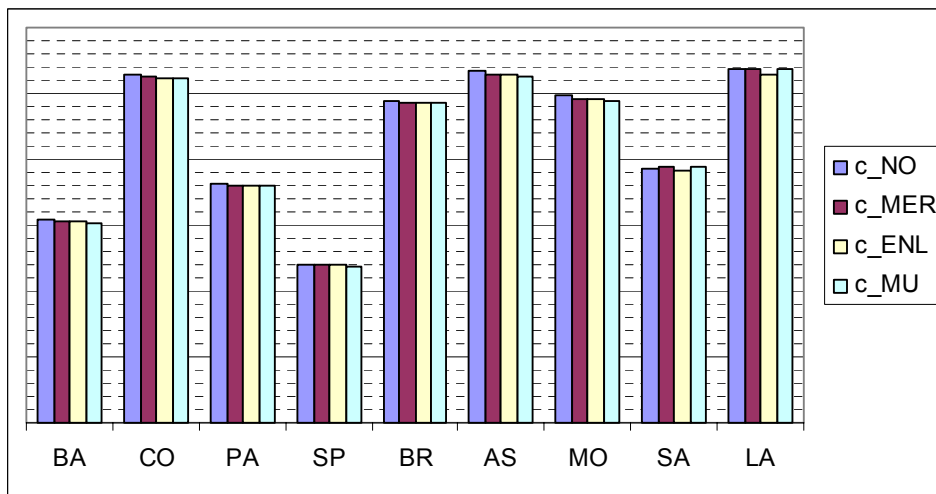
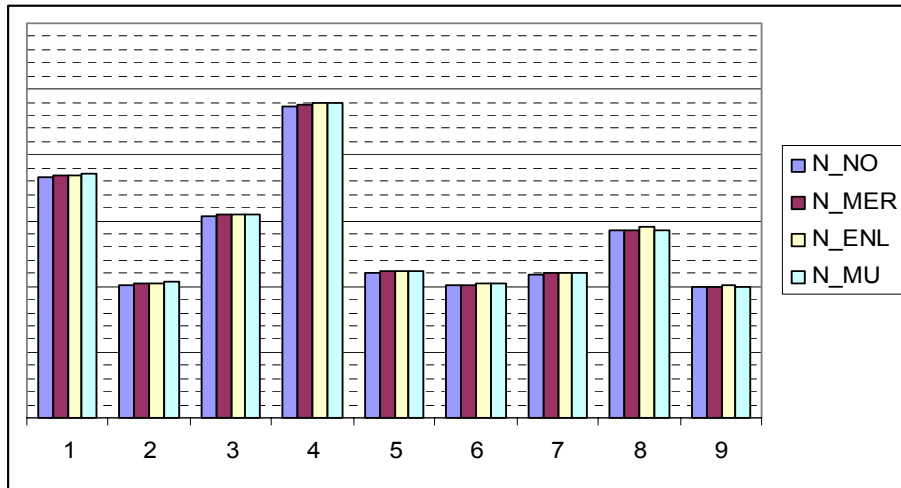


Figure 9 – Regional integration: Cost-saving attraction and accessibility



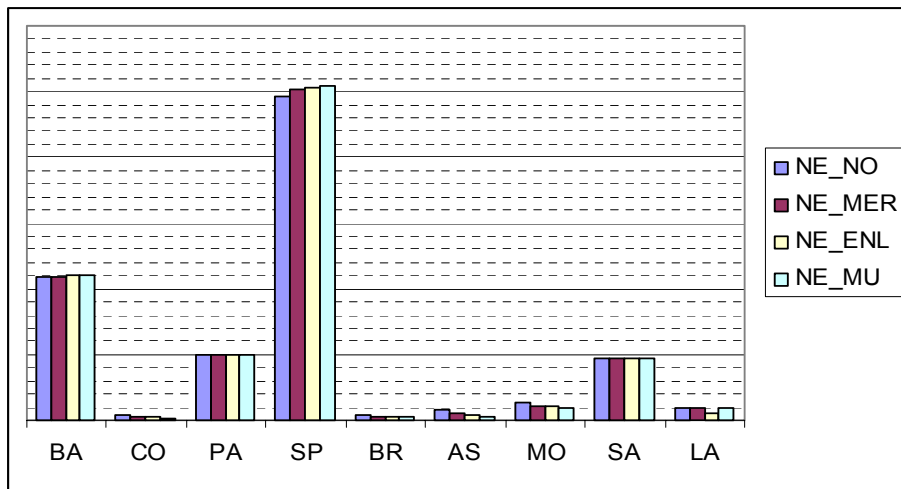
Scenarios: No trade liberalization (NO), Preferential trade liberalization excluding LA and SA (MER), Enlarged trade liberalization (ENL), Preferential trade liberalization plus monetary union excluding LA and SA (MU).

Figure 10 – Regional integration: Selection



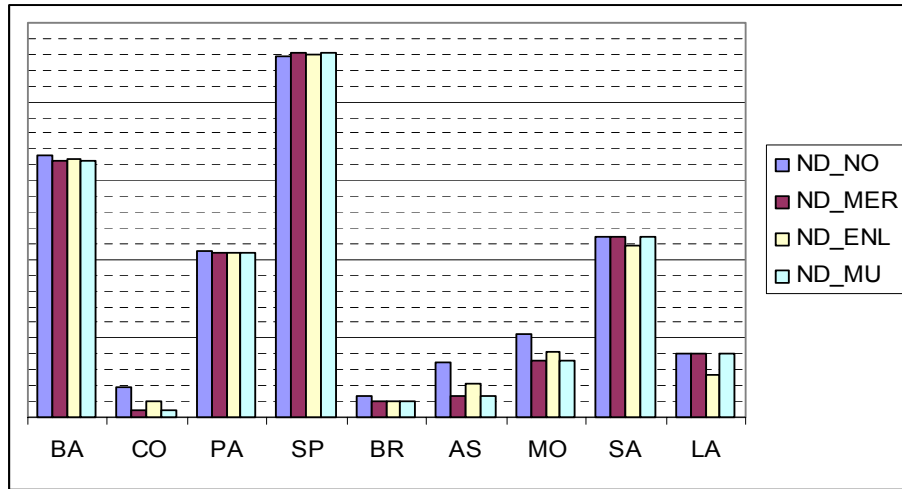
Scenarios: No trade liberalization (NO), Preferential trade liberalization excluding LA and SA (MER), Enlarged trade liberalization (ENL), Preferential trade liberalization plus monetary union excluding LA and SA (MU).

Figure 11 – Regional integration: Variety



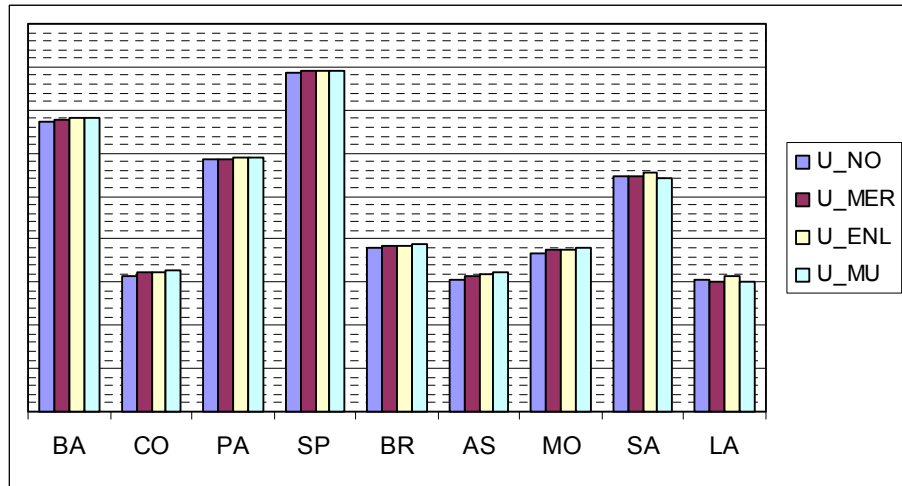
Scenarios: No trade liberalization (NO), Preferential trade liberalization excluding LA and SA (MER), Enlarged trade liberalization (ENL), Preferential trade liberalization plus monetary union excluding LA and SA (MU).

Figure 12 – Regional integration: Entry



Scenarios: No trade liberalization (NO), Preferential trade liberalization excluding LA and SA (MER), Enlarged trade liberalization (ENL), Preferential trade liberalization plus monetary union excluding LA and SA (MU).

Figure 13 – Regional integration: Production



Scenarios: No trade liberalization (NO), Preferential trade liberalization excluding LA and SA (MER), Enlarged trade liberalization (ENL), Preferential trade liberalization plus monetary union excluding LA and SA (MU).

Figure 14 – Regional integration: Welfare