1. Introduction

Economic activities are unevenly distributed across space. The determinants of spatial differences in the patterns of production have traditionally been presented in terms of differences in endowments, technologies, market scheme, or policy regimes.
Such explanations, while relevant, fail to explain why even a priori similar regions (or countries) can develop very different production structures. They also stand in sharp contrast with the changing pattern of comparative advantage of regions and countries undergoing rapid development.
Recent contributions (1990- now) to the general equilibrium of trade theory associated with industrial organization, labor economies, and regional economies have developed a novel approach to the way we think about location; one in which firms tend to cluster together, and regions with similar –or even identical- underlying characteristics can turn out to be very different.
Much of that work focuses on how the propensity of firms and workers to agglomerate in space changes as regions become more integrated.

**Industrial agglomeration** can be considered at different levels of aggregation. Starting from the **bottom**, there are small scale agglomerations of finely defined sectors.
Amongst the best known examples of such highly specialized industrial districts are the US carpet production industry in the Georgian city of Dalton (Krugman, 1991) and the Italian textile industry in the city of Prato (Pyke, Becattini and Sengenberger, 1990; and Porter, 1900).
At the other end, there are **large scale agglomerations** that cut across state and country boundaries.

These include the US ‘Manufacturing Belt’ (approximately contained in the parallelogram Green Bay-Saint Louis-Baltimore-Portland) and the European ‘Hot Banana’ (the area between Milan and London, containing Northern Italy, Southern Germany, South East France, Belgium, the Netherlands, and South East England).
The possible explanations of why firms tend to cluster together, *technological externalities* arising from personal interactions matter (*R&D*) most for small scale agglomerations.


We are also concerned with understanding large scale agglomerations, creating differences between entire regions.

➤ What is at the origin of such inequalities?

➤ How will the spatial distribution of economic activities change as countries and regions become more integrated?
Will economic integration lead to a higher or lower degree of spatial agglomeration?

And are all activities likely to evolve similarly, or will integration have a different impact on different sectors?

Addressing those questions with a common set of tools, involving a combination of monopolistic competition à la Spence (1976, RES) and Dixit and Stiglitz (1977, AER) and ‘iceberg’ trade costs à la Samuelson (1954, Economic Journal).
These tools, while peculiar, are the best available for dealing with increasing returns, trade costs, migration, and input-output linkages between firms in an analytically tractable general equilibrium framework.

An excellent recent survey of how this approach relates to work based on either technological externalities or spatial competition is provided by Fujita and Thisse (Economics of Agglomeration, Cambridge, 2002).
2. The viewpoint of traditional and ‘New’ trade theory

*Comparative Advantage*

- *Economic theory* has traditionally explained differences in production structures mainly through *differences* in underlying characteristics (*geography, endowments, technology, policy*), which make space itself uneven.

- In this framework, *industrial integration* leads regions to specialize according to their comparative advantage (see, e.g., Jones, 1965).
Absent underlying differences between regions, models of trade with **constant returns to scale** and **perfect competition** predict that economic activities will be evenly distributed across space.

Product and factor market competition provide reasons for firms to locate far from each other, but these must be set against those forces which tend to pull firms together.
Comparative advantage, while relevant, provides a weak explanation for the remarkable spatial concentration of activity—often very similar regions have very different production structures.

Furthermore, when more things are mobile than not, especially on the globalization nowadays, traditional trade theory fails to give clear-cut predictions of the patterns of specialization and trade.
2.1 Market Accessibility

Increasing returns to scale turns out to be essential for explaining the uneven geographical distribution of economic activity.

Models of trade with increasing returns and imperfect competition explain why countries without significant comparative advantage with respect to each other can develop different production structures on the basis of their different access to markets.
2.2 Endogenous Agglomeration Patterns

- Models of trade with imperfect competition predict that, in the presence of increasing returns and trade costs, firms and workers tend to locate close to large markets.
Development economists in the 1950s were keen to emphasis that large markets are in turn those where more firms and workers locate.

New trade theory (or economic geography) formalizes this kind of cumulative causation mechanisms, to show that countries (or regions) which are similar, or even identical, in underlying structure can endogenously differentiate into rich ‘core’ countries and poor ‘peripheral’ countries.
There are several mechanisms through which such cumulative causation may arise.

Krugman (1991, *JPE*) shows that the interaction of labor migration across regions with increasing returns and trade costs created a tendency for firms and workers to cluster together as regions integrate.

While relevant for studying agglomeration within national boundaries, in an international context high barriers to migration may limit the role of labor mobility as a force driving agglomeration.
Figure 1. The centripetal force of spatial economy
Venables (1996, IER) addresses this issue by showing that vertical linkages between upstream and downstream industries, when both of them are imperfectly competitive, can play a role equivalent to that of labor migration in endogenously determining the size of the market in different regions.
2.3 Migration-induced Demand Linkages

→ The main point of Krugman (1991, *JPE*) is that, if some factors are mobile between countries, then the pressure put on those factors by the concentration of economic activities will be eased.

→ Factor mobility will make the supply of factors sufficiently elastic that small differences in the size of industry across regions can build up.

→ Even if regions are a priori identical, they can become endogenously differentiated into an industrialized core and a deindustrialized periphery.
Two countries (or regions) and two sectors, one monopolistically competitive (industry) the other perfectly competitive (agriculture), and two production factors.

First, each of the two sectors uses a specific factor, so there is no intersectoral reallocation of factors (as we discuss below, this element is not essential to the story).

Second, only the factor used by industry (workers) is mobile between regions, while the factor (farmers) used by agriculture (farmers in Krugman’s paper, although we find it useful to think of this factor as land) is not.

Finally, the two regions are a priori identical in every respect, including their endowment of immobile factors.
Suppose that for some reasons one firm decides to move production from one region (or country) to the other.

How does this affect firms’ profitability?

The presence of one more firm will *increase* competition in the product and labor markets of the region receiving the firms, thus tending to reduce local profits and to make relocation unprofitable.
➔ If there was **no migration**, this would be the end of the story and regions would remain identical.

➔ However, the rise in the number of local varieties and the rise in labor demand and wages tend to **attract** more workers.

➔ This increases local expenditure (a **demand linkage**) and eases competition in the labor market, and so tends to increase local profits and to attract more firms.
Whether the overall effect of entry is to increase the profitability of local firms (encouraging further entry), or to lower that profitability (leading to exit), depends on parameters of the model, and in particular on how integrated regions are.
Agglomeration takes place earlier during a process of regional integration the stronger the preference for variety, and the larger the share of manufactures in expenditure.

This is because a lower elasticity of substitution across varieties in consumers’ preferences increases the importance of having a large variety of products available locally.
By reinforcing the monopoly power of firms over their own varieties, this *weakens* local competition and *favors* agglomeration.

A larger share of manufactures in consumer expenditures also favors agglomeration, because it augments the impact of immigration on the size of the local market for manufactures.

It increases the weight of the prices of manufactures in real wages, thus enabling firms located in regions with more industry to attract workers without having to pay high nominal wages.
2.4. *Input-output Cost and Demand Linkages*

The cumulative causation mechanism modeled by Krugman (1991, *JPE*) relies on the assumption that, when a country does relatively badly in terms of non-agricultural employment, workers move to country that are doing relatively better, and this tends to eliminate international (or interregional) real wage differentials.
Venables (1996, *IER*) addresses this issue by arguing that firms like to be close to each other not only because of linkages working through the supply of labor and demand for goods from each other’s workers, but also because of direct *input-output linkages* amongst themselves.

This amounts to a formalization of Hirschman-type (1958) ‘*forward*’ and ‘*backward*’ linkages between industrial firms in the economy.
For a **downstream industry** to bestow a **backward linkages** on an **upstream industry**, it is not enough that there is a buyer-supplier relationship between the two; it must be the case that an increase in the output of the downstream industry, by enlarging the market for the intermediates it uses, induces the upstream industry to produce at a more efficient scale.
Similarly, a **downstream industry** enjoys a **forward linkage** only in so far as an increase in the output of an upstream sector allows downstream firms to produce more efficiently.

Recently, **upstream-downstream industry** associated with the issue of industrial organization and growth:

Venables (1996, *IER*) considers an economy with two countries and internationally immobile labor. Besides a perfectly competitive sector, he models an upstream and a downstream imperfectly competitive industries, where the goods produced by upstream firms are inputs to downstream firms.
The structure of the model by Venables (1996, *IER*)

Figure 3. The structure of the model
Krugman and Venables (1995, *QJE*) make the model closer in structure to Krugman (1991, *JPE*), by collapsing the upstream and downstream industries to a single imperfectly competitive sector in which the output of each firm is sold both as a final good to consumers and as an intermediate input to all other firms.
4. **The structure of the model** Krugman and Venables (1995, *QJE*)

**Agricultural Sector**
- Constant Return to Scale
- Labor input

**Manufactured Sector**
- Monopolistic Competition
- Labor input
- Import the manufactured goods from the other Regions as intermediated inputs
- Intermediated inputs in the same region
- Export the manufactured goods to the other region as intermediated inputs

- Cobb-Douglas Utility Function
- Final consumption
- Manufactured Goods
- Agricultural Goods
- Export the manufactured goods to the other region as intermediated inputs
- Import the manufactured goods from the other Regions as intermediated inputs
- Intermediated inputs in the same region
- Final consumption
The linkages driving **agglomeration** are different.

In Krugman’s (1991, *JPE*) model, an increase in the number of firms in a location increases demand for the output of local firms through the expenditure of the workers attracted from other regions.

In Krugman and Venalbes (1995, *QJE*) there is no interregional mobility, so workers must be drawn from other sectors instead, and the higher demand comes from expenditure on intermediated by the newly arrived firms.

In Venables (1996, *IER*) there is a cost linkage arising from the saving in trade costs on a larger fraction of their intermediate inputs by firms in the larger market.
2.5. *Endogenous Growth, Factor Accumulation and Intertemporal Linkages*

Baldwin (1997) suggests an alternative way in which agglomeration may occur without factor migration. He shows that factor accumulation can play the same role as migration in fostering agglomeration through demand linkages.
The structure of his model is similar to that of Krugman and Venables (1990) with the addition of a Research and Development (R&D) activity that uses labor to invent and patent new manufactures.

The differentiated manufactures are produced using labor and patents as an additional factor. Patents last forever and are non-tradable, so that production occurs where invention takes place.
What is then the spatial distribution of firms at this equilibrium?

The answer depends on whether profits accruing to new patents increase or decrease with the number of firms.
These profits are determined by the balance between the attractiveness of the larger market in the presence of increasing returns and its unattractiveness arising from product and factor market competition.

If the former dominates, invention in the larger market pays more and endogenous accumulation prevents an even distribution of firms.

The way in which the model generates *circular causation* between concentration and growth can be seen by considering the usual firm relocation starting from an even spatial distribution of firms and labs.
The presence of one more firm gives labs access to a wider range of services without any additional trade cost. This makes innovation cheaper.

Local labs innovate at a faster rate and some labs relocate from the other country.

Faster innovation and more labs in turn increase the local demand for intermediates and therefore attract more industrial firms. So labs follow firms and firms follow labs.
3. Globalization and the spread of industry

3.1. Labor Immobility as a Dispersion Force

Puga and Venables (1998, *JIE*)

The model considers both interregional migration and input-output linkages as forces which may drive agglomeration; it also takes a closer look at the interaction between constant and increasing returns activities in labor markets.
Four main conclusions can be drawn from that analysis.

➢ **First**, comparison of the outcomes with and without interregional migration shows that agglomeration gets an extra kick from the relocation of workers towards locations with higher real wages. Thus the lack of *interregional mobility* both postpones agglomeration in a process of regional integration and weakens it when it happens.
Second, if equilibrium wage differences are not eliminated by migration, they act as a dispersion force by increasing production costs for firms producing in locations with relatively many other firms.

Third, this dispersion force can moderate agglomeration and sustain non-extreme equilibria in which all regions have industry, even if in different proportions.

Fourth, firms find higher local wages increasingly discouraging as regions become more integrated, so for low trade costs it is the price of non-tradable factors that determines location.
3.2. *Non-tradeability as a Dispersion Force*

- Wage differences are one possible reason why globalization can bring convergence of income levels, but not the only one. An interesting alternative is modeled by Helpman (1997).

- He takes Krugman’s (1991, *JPE*) model and turns the agricultural sector producing a freely tradeable commodity into a **non-tradable** housing sector.
Helpman finds the reverse results to those in Krugman’s paper: reductions in trade costs improve the availability of manufactures in less congested areas and induces workers to migrate out to more congested areas to save on housing costs, thus working against agglomeration.

Interestingly, intermediate cases between Krugman’s and Helpman’s models (studied by Adrian, 1996; and by Hadar, 1996) produce a picture in which industry is agglomerated for intermediate values of trade costs, but not for high and low values of trade costs.
The general picture coming out of these models is therefore one in which, for high trade costs, the need to supply markets locally encourages firms to located in different countries (or regions).

For intermediate values of trade costs, cost and demand linkages take over and firms and workers cluster together.

Finally, for low values of trade costs, firm’s location is determined by the price of those factors and goods that are not mobile.
3.3. **Industrial Specialization**

- Krugman and Venables (1996, *EER*) show that the observation that firms have stronger buyer/supplier relationships with some types of firms that with others can help us understand the process of regional specialization.
They consider a setup like that in Krugman and Venables (1995, *QJE*), with one main difference:

- the two production sectors are imperfectly competitive, and firms in each sector sell and buy a higher proportion of intermediates to and from firms in the same sector than to and from firms in the other sector. (i.e., Trade: Intra-industry > Inter-industry)

- The **forward** and **backward** linkages operating in this case are essentially the same as in Venables (1996, *IER*), Krugman and Venables (1995, *QJE*).
The difference is that, if one more firm locates in a country (or region), the benefits of cost and demand linkages affect more intensely firms in the same sector, while the increased product and labor market competition harms firms in both sectors equally.

As a result, integration leads each country (or region) to become specialized in the production of one sector.
4. Concluding remarks

- Firms producing in locations with relatively many firms face stronger competition in the local product and factor markets.

- This tends to make activities dispersed in space.

- However, the combination of increasing returns to scale and trade costs encourages firms to locate close to large markets, which in turn are those with relatively many firms.

- This creates externalities which favor the agglomeration of economic activities
➢ **Economic integration**, by affecting the balance between *dispersion* and *agglomeration* forces can decisively affect the spatial location of economic activities.

➢ For **high trade costs**, the need to supply markets locally encourages firms to locate in different countries (or regions).

➢ For **intermediate values of trade costs**, the incentives for self-sufficiency weaken.
➢ **Externalities** then take over, and firms and workers cluster together.

➢ However, the **price of local factors and goods** tends to rise wherever agglomeration takes place.

➢ If most factors and goods can be imported from other countries, rising factor prices simply give an additional kick to agglomeration by inducing immigration.
If instead there are some **immobile factors** which are particularly important for production (such as labor), or non-tradable goods that are particularly important for consumption (such as **housing**), as further integration reduces the importance of externalities, differences in the prices of immobile goods and factors take over.
What arises for the welfare (or indirect utility) inequality between developed country and developing country in a stage of globalization?
Figure 5: Welfare (or indirect utility) inequality

Source: Krugman and Venables (1995, *QJE*)