Trade and Spatial Growth: the Nexus that was not missing

Sirimal Abeyratne
Visiting Research Fellow, IRI Research Institute, International University of Japan

N.S. Cooray
International University of Japan

April 2014
Trade and Spatial Growth: the Nexus that was not missing

Sirimal Abeyratne
University of Colombo

and

N.S. Cooray
International University of Japan

ABSTRACT

Trade leads to an acceleration of economic growth as well as to its spatial concentration. While trade-growth nexus has been the primary focus in trade analyses, until the recent past spatial growth concentration has received little space there. It appears that the simplifying assumptions of trade models and analyses have become the main stumbling blocks that push spatial growth away from trade theory. It is the rapid change in the shape of the world economy led by trade liberalization and global integration that have created this space to place intra-national economics within the premises of international economics.

Spatial growth within trade analyses was, however, not bizarre to trade theory. Bertil Ohlin as well as his teacher, Eli Heckscher – the coauthors of the Heckscher-Ohlin theory of comparative advantage – did not miss the point that comparative advantage is based on location-specific factors. The analyses of the locations of production and economic geography seem to have missed the early contribution to the subject by Ohlin and Heckscher who were not confined themselves by the simplifying assumptions of their own trade theory. The paper derives from its literature review that the benefits of agglomeration, the costs of connectivity, the degree of factor mobility, and the size of the markets are essential components that form centripetal forces of spatial growth concentration.

JEL Classification: F12, F13, F15, F16, R11, R12

Key terms: Economic geography, International trade, Regional development, Spatial economy

The Working Paper is based on the Research Project on Trade and Spatial Growth: Sharing Images from Japan and Sri Lanka, carried out by the authors in 2013 at the International University of Japan under the Fellowship granted by the Japan Foundation. The authors wish to thank both institutions and many others who have contributed to the research project in numerous ways.
1 Introduction

Trade leads to an acceleration of economic growth as well as its spatial concentration. The link between trade and spatial growth was not bizarre to Bertil Ohlin, who seemed to have been inspired by his teacher Eli Heckscher – the coauthors of the Heckscher-Ohlin theory of comparative advantage. Ohlin, who titled his publication in 1933 as *Interregional and International Trade*, treated trade theory as part of location theory (Ohlin, 1967). Eli Heckscher, in his Swedish publication in 1919 analyzed the distributive implications of Ricardian assumption of factor immobility (Heckscher, 1950). However, the initial contributions of Ohlin and Heckscher to the analysis of the location of production have almost been forgotten now as there has been no fair acknowledgement to these contributions in the relevant studies. Apart from that until the recent past international trade and spatial growth appear to have developed as two distinct branches of Economics. The former was occupied overwhelmingly by the analyses on trade-growth nexus, leaving its spatial dimension issue behind. The later was occupied by the analyses of regional inequality which received topical interest of diverse perspectives within development economics, apart from its overwhelming occupation in regional and urban studies as well as in economic geography.

The purpose of this paper is to provide a review of the theoretical premises of the link between international trade and spatial growth, and thereby to outline the basic factors underlying spatial growth concentration. The paper is intended to link current discussion on trade and spatial growth to its initial conceptualization within the premises of trade theory with a brief account of its evolution.

The paper is organized as follows: Next section presents a review on growth acceleration and concentration as interrelated economic phenomena. This is followed by an analysis of spatial growth in the context of Heckscher-Ohlin trade theory and the contributions to it made by Bertil Ohlin and others. As it is briefly outlined in the next section, however, spatial growth received attention in development theory from diverse perspectives. As discussed in the subsequent sections, while the assumptions of trade models which became conditional in most of the subsequent trade analyses led to an abandonment of spatial growth analysis within trade theory, this was challenged by the reshaping of the world economy in the past few decades creating space for revisiting the missing dimensions. Finally, the factors underlying trade and spatial growth as evolved from literature review are presented. The paper concludes that the nexus between trade and spatial growth was not missing in initial trade analysis though it was missed subsequently.

2 Growth: Acceleration and Concentration

As countries grow, economic activity and people concentrate on certain specific locations creating ‘leading’ regions as metropolitan cities, urbanized areas, commercial hubs, and industrial clusters. These specific locations with greater economic and population densities play a major role in both national and global economic spheres, leaving larger regions as sparsely-populated rural farmlands. Spatial concentration of economic activity and people can be observed in both developed and developing
countries, but economic growth speeds up the process of spatial concentration. Naturally countries continue to develop with the growth of their cities, while these cities become leading areas that make influential contributions to national output growth. World Bank (2009:48) states that ‘as countries become richer, economic activity becomes more densely packed into towns, cities, and metropolises’. Geographical scale does not matter; just as the regions of the world such as Western Europe, North America, and East Asia differ from the rest of the world in terms of economic density according to the global scale, growth results in unevenness in spatial growth within the countries and even in smaller countries and regions.

Geographical attributes, historical events, political decisions, and even accidents may have been important in the choice of locations as leading regions at initial stages, but the process of spatial concentration of economic activity and people in these locations is directly linked to economic growth. Rapid growth increases the gap between leading and lagging regions, as production continues to concentrate more in leading than in lagging regions. Once these leading regions enter into a self-sustaining stage, the initial factors of locational advantage become less and less important, and the multiplier effects of spatial concentration become more and more important.

In every country – developed or developing, there are main cities and metropolitan areas which play the most important role in their national economy in making a disproportionately high contribution to GDP. Generally, these locations in the respective countries appear as commercial and financial hubs, industrial agglomerations, political and administrative centers, and international gateways through rivers, seaports and airports. Even in smaller geographical units of a country such as remote regions, which do not possess these strategic features of national importance, business activities get concentrated in small townships. In large countries such as USA, China, and India as well as in smaller countries such as Netherlands, Belgium, and Sri Lanka, economic activity and people are concentrated in specific regions, cities and urbanized areas. In developed countries in the Western Europe and North America as well as in developing countries in Africa, Asia, and Latin America, there is spatial concentration of economic activity and people in urbanized areas, cities and metropolises.

There are centripetal forces in attracting economic activities closer to each other in specific locations, in contrast to those centrifugal forces which push them away from such locations. There are both forward and backward linkages of economic activities which pull them together in specific locations against those forces which push them apart in remote areas. Economic activities find it advantages for them to locate themselves close to their output markets (forward linkages) and input markets (backward linkages). It is even more attractive for economic activities to stay close to the bigger markets than to smaller markets so that metropolises generate stronger centripetal forces than the smaller cities. For this reason, it is not unusual to observe that metropolises become even more concentrated, and grow even faster than the smaller cities.

Producers also find it important for them to stay close to each other even if they are rival firms producing identical or differentiated commodities focusing on the same markets. The centrifugal forces naturally lead the rival firms to locate far from each other by inhibiting positive externalities, closing information exposure, avoiding competition, and protecting the market share. In spite of all these
benefits, the centripetal forces that keep them together are much stronger than centrifugal forces that fall them apart.

Spatial concentration is not limited to the formation of industrial clusters – the location of firms operating within the same industry. Although there is a degree of specialization, divergent and un-related economic activities in a wide range could be seen concentrating in cities, metropolises, and industrial agglomerations. Although they are highly divergent and un-related, they all derive the benefits of being together.

Growth does not lead the same types of economic activities to grow and get concentrated over geographical space. As different economic activities grow differently, spatial growth is accompanied by structural transformation of economic activity. Even at an aggregated level, it quite apparent that manufacturing and service activities in a country are concentrated in one or few smaller locations, leaving large stretches of land as farms and forests. In the USA and the Western Europe where large-scale agriculture production exists, there is a clear distinction between sparsely populated large areas of farm belts and highly congested metropolitan cities and industrial agglomerations. Even in countries such as Japan where traditional institutions and regulatory barriers have blocked the formation of large-scale farm belts, the structure is not contrary to that, though magnitudes are different. Highly congested industrial agglomerations and metropolitan cities have emerged and continued to grow leaving large parts of the country-side as sparsely populated rural villages and forests.

The structural changes accompanied by economic growth have followed a pattern in the process of moving from an agriculture-based economy towards a modern economy dominated by industry and service sectors. As a result, the share of agriculture output has declined to around 1 percent of GDP, and the share of employment in agriculture sector to less than 5 percent of the workforce in most of the advanced countries, allowing the rest of the output and employment to be replaced by growing industrial and service sectors. Even within the industrialization process, countries have begun with producing labor-intensive light manufactures, moving into capital-intensive heavy industries, and then to advanced types of manufactured production dominated by labor-saving, energy-saving, knowledge-based, and high-tech sectors. The process depicts the changes in comparative advantage as was observed and documented in the studies of emerging economies in Asia and elsewhere following policy reforms since the 1970s. Similarly, financial sector reforms in many countries have led them to grow and emerge as regional or global financial hubs which had implications on their output and employment structures.

The seminal work of Williamson (1965) confirms the existence of regional inequality in the development process of a number of developed and developing countries. The study resembled the Kuznets-type inverted U-shaped regional inequality curve associated with economic growth; the study of Kuznets (1955) revealed a rise and fall in income inequality respectively at the early and latter stages of growth. As the study of Williamson (1965) concludes, there is increasing regional inequality at the early stages of growth, while matured growth in developed countries has produced a regional convergence or a reduction in regional inequality.
The economic disparity between leading and lagging regions falls at latter stages of growth, because of the spillover effects of growth on both production and consumption in lagging regions. As the countries become richer, they have the capacity to connect lagging regions to the leading ones, to transform the production patterns in the lagging regions, and to divert the benefits of higher growth from leading to lagging regions.

Regional differences in per capita GDP within the countries are substantial and larger in the rest of the world than among OECD countries, indicating that the level of development has implications on spatial growth and regional income disparities. According to OECD (2011:41) estimates of the Gini Index of inequality, the regional inequalities of per capita GDP are much larger in Russia, India, China, and Brazil than their OECD average. The evidence shows that the regional disparities resulting from spatial concentration of economic growth is a common phenomenon, and they tend to widen as the countries grow. As spatial growth concentration is accompanied by structural changes and demographic transition, the regional disparities in per capita income decline at mature stages of growth.

The demographic transition with internal migration contributes to the divergence of regional inequality at the early stages of growth and to the convergence of it at the latter stages of growth. The spatial growth pulls population from lagging areas to leading areas. Leading regions provide better opportunities for human resources to be utilized more productively in the growth process as well as to derive the benefits of growth in terms of higher living standards. A detailed analysis of internal migration associated with spatial concentration of economic activity exhibits different patterns of human resource development and specialization of economic activity. Therefore, it is also observed that skilled labor and educated people concentrate more in line with spatial concentration of economic activity, as they find more opportunities and tend to benefit more from their concentration in leading areas than in lagging areas.

The poor also concentrates more in leading areas than in lagging areas. This leads to poverty implications of spatial concentration of economic activity in a country. As World Bank (2009) also noted, lagging regions account for a greater share of poor because economic backwardness itself is the cause of poverty; but more of the poor appear to have concentrated in leading regions. Because the lagging regions are relatively less productive as denoted by the lower economic density, generally the share of poor is greater in lagging regions than in leading regions of a country. As people get concentrated in line with spatial growth, however, many poor also get concentrated in the leading than in the lagging regions. This is because many poor find it more advantageous to live in leading areas where they can also find more opportunities than in lagging areas in order to escape from poverty.

3 From International to Intra-national Economics

Traditional trade theory, since its early stages of development, is concerned with ‘location-specific’ factors as the determinants of comparative advantage of the location, showing that the two branches of economics – international trade and spatial growth, emanate from the same source. This idea was in the heart of trade theory so that Bertil Ohlin who considered international trade theory as part of the
‘location theory’, titled his book first published in 1933 as *Interregional and International Trade*. His teacher, Eli Heckscher in his 1919 publication was aware of the implications of the assumption of ‘factor immobility’ in the Ricardian theory of comparative advantage.\(^1\) Although Heckscher-Ohlin theory of comparative advantage became a dominant analytical tool in trade analysis, it was trade-growth nexus, and not the ‘location of production and specialization’ that received much attention in subsequent trade analyses.

Economic modeling on spatial growth began as far back as 200 years ago when Johann Heinrich von Thünen (1783-1850) who in his publication of the English translation in 1826, *The Isolated State*, modeled the concentration of economic activity in the city surrounded by agricultural hinterland, and thereby the determination of wages and rents (Samuelson 1983). His model of spatial growth was based on that labour is mobile, but land is immobile, and that there is transport cost to move goods across space – that is, between the city and the hinterland in an ‘isolated state’. Interestingly, according to Samuelson (1983), Thünen worked out his model determining wages and rents even before the work of the contemporary classical economist, David Ricardo (1772-1823), and used the exact opposite assumptions of Ricardian trade theory – factor immobility and zero transport costs. Thünen’s work, in spite of its relevance and importance as an early contribution to the location of production was not necessarily within the context of trade theory.

The Ricardian theory established the cross-country differences in relative labour productivity or, generally the comparative cost advantage in producing different commodities as the basis for trade and gains from trade. Hecksher-Ohlin theory that disclosed a new dimension of comparative advantage was founded upon the cross-country differences in factor endowment. The standard assumptions of the theory include, among other things, the immobility of productive factors and their fixed supply, the absence of transport and information costs, and the existence of linearly homogeneous production functions resulting in constant returns to scale. Thus the Hecksher-Ohlin theory predicts that under free trade condition countries gain from trade by specializing and producing the commodities which intensively utilize its relatively abundant factors, and exchange them (export) for other commodities (import) which are intensively utilized its relatively scare factors. In spite of the fact that the Heckscher-Ohlin model is based on simplifying assumptions, its co-authors – Eli Heckscher and Bertil Ohlin, did not adhere to such assumptions in their trade analyses which led them to analyze trade implications on a borderless geographical space.

Eli Heckscher, in his paper on ‘The Effect of Foreign Trade on the Distribution of Income’, published in 1919 categorically referred to the Ricardian assumption of ‘complete immobility of factors’, and Edgeworth’s account that ‘international trade means exchange on the basis of immobile factors of productions’ (Heckscher 1950:285). This implies, in other words, that trade without borders cannot exist, or that the existence of trade within borders needs to be explained in a different context. However, Heckscher (1950: 277) noted that ‘if Ricardo’s assumption of immobility of factors of production

---

\(^1\) It has been a debatable issue whether the theory of comparative cost advantage was discovered by David Ricardo (1772-1823) or by Robert Torrens (1780-1864), who wrote on international trade prior to Ricardo. Neither of them used the term ‘comparative cost’, which was actually done by James Mill (1773-1836); for a recent synthesis on this debate, see Ruffin (2002).
between countries is dropped, it can easily be seen that the determination of factor prices has an important bearing upon factor movements’. As the assumption of the immobility of productive factors is only partially valid for labour and capital both, though it is not so for land, Heckscher’s analysis was concerned with trade effects on factor prices under different scenarios of factor mobility. If there is full mobility of productive factors, apart from transportation costs, factor prices tend to be equalized and trade tends to cease, which is an outcome that is beyond the scope of Ricardo’s model.

Bertil Ohlin who considered interregional trade as equally important as international trade, disclosed the interrelationship between the location theory and international trade in his book *Interregional and International Trade*, first published in 1933 (Ohlin 1967). According to Ohlin (1967: 97), ‘...the important distinction is not between domestic and international trade theories, but between a one-market and many-market theory of pricing’. Ohlin did not rule out the assumption of the immobility of productive factors. There is no doubt that Ohlin would have been inspired greatly by the work of his teacher Eli Heckscher, who in his subsequent writings appreciated the extensive work of the student in the area of factor immobility (Heckscher 1950). Ohlin, who made extensive observations on the factors underlying factor immobility assumption of trade theory, did consider that distance between the locations matter so that exchange is costly whether it is international or domestic trade. This approach essentially removes some of the important barriers within trade theory to the analysis of the location of production and exchange irrespective of the geographical scales:

It is true that in a study of location in international trade the lack of factor mobility is perhaps the most important element (although there are also special obstacles to international commodity movement that must be reckoned with); but *international trade theory cannot be understood except in relation to and as part of the general location theory, to which the lack of mobility of goods and factors has equal relevance* (Ohlin 1967: 97).

By the clause *the lack of mobility of goods and factors* which has equal relevance to trade and location, Ohlin stressed trade and locational implications of the assumptions of both factor mobility and transport cost. Generally, natural resources (land) are immobile, but there are non-economic obstacles to labour and capital mobility which might be overcome through an additional economic cost – the payment of higher factor prices. Thus, according to Ohlin, the degree of factor mobility has a bearing impact on factor prices and commodity prices as well as on trade. Ohlin (1967: 26) did not, however, anticipate a complete equality of factor prices, which he considered as ‘almost unthinkable and certainly highly improbable’. He asserted that industry demand for factors of production is always a ‘joint demand’ for several factors. As a result even within a country ‘a set of factors’ is not completely mobile so that Ohlin justified the existence of obstacles to free factor mobility. This postulation implies that productive factors are not homogeneous aggregates as they assumed to be as labour, capital or land, but there exist within each category a wide variation of sets of factors.

Ohlin incorporated *economies of scale* emanating from the regional concentration of labour and capital, perhaps for the first time in trade analysis, as an important determinant of export expansion:
...The large-scale economies effected when the market is enlarged through the influx of labor and capital affect the various industries quite differently. If these economies are felt chiefly in export industries, trade tends to increase. However, increased effectiveness, particularly in industries that produce goods competing with import commodities, must have the opposite effect on the volume of trade. (Ohlin 1967: 120)

He extended the discussion on the benefits of the large scale economies – both internal and external to the firms, to cover the transport sector, thereby asserting the importance of the linkages between transport services and the industries producing commodities. The variations in factor supply and factor prices also change the costs of transfer as transport services depend, in turn, on the same factor prices as well as the economies of scale. Therefore, according to Ohlin (1967: 121), ‘harbors, and railroads, for example, may be built and cheaply operated, which must markedly affect the volume and character of trade’. Apart from the locational implications of internal and external economies of scale, by these examples of harbors and railroads, Ohlin showed the importance of the presence of non-tradable goods as important factors that keep pulling industries to specific geographical locations.

Ohlin has employed the basic premises of the Heckscher-Ohlin theory to analyze trade without geographical scales and the location of production and specialization. While qualifying and maintaining the assumption of ‘imperfect’ factor mobility, he has brought about many of the basic elements of the link between trade and spatial growth into the analysis; these aspects included the transport costs, geographical concentration of productive factors, presence of economies of scale, and the location of input and output markets.

The theoretical premises of international trade has provided with scope for analyzing the geographical location of production and trade. In the words of Higgins and Savoie (1988:9), ‘...the trade approach, when applied regionally, would see regions specialize in their areas of strengths and comparative advantage’. For instance, Couchene and Melvin (1988) presented a standard neoclassical trade model in the context of Hecscher-Ohlin theory in order to explain regional growth and income disparities within two regions of a country, where trade exist both domestically between the two regions, and internationally with the rest of the world. A criticism of growth pole theorists on standard neoclassical trade model was that it has failed to offer a theoretical application to both international trade and domestic trade. In response to this criticism, Couchene and Melvin (1988) attempted in their trade model to show that neoclassical theoretical premises has adequate space to analyze regional locations of economic activity within a country which is open to free trade within the regions of a country as well as between its regions and the rest of the world.

As it is argued in this publication too, the issue in question was centered on the strict assumptions of neoclassical trade model which limits its application to domestic trade. The standard trade model is based on the production possibility frontier, terms of trade, and community indifference curves which are constructed on the basis of standard neoclassical assumptions. The analysis is simply an extension of the neoclassical model on trade between two countries to one between one country with two regions and the rest of the world; by implication there is domestic trade between the two regions, and international trade with the rest of the world. This extension required a modification of the standard
assumptions of the two-country trade model to suit the extended trade model of two-countries and, one of them with two regions.

Under the different scenarios related to the assumptions of resource endowment, factor mobility, transport costs, and consumer preferences, the paper is aimed at modeling interregional income disparities. According to the authors, ‘...this extension has implications both for the standard theory of international trade and for regional economic policy analysis’ (Couchene and Melvin 1988: 171). As far as spatial growth is concerned the policy relevance of the extension is limited, but it clearly shows that the interregional trade and its implications on regional income disparities can be analyzed within standard trade models by modifying the underlying assumptions.

The strict adherence to the simplifying assumptions of trade models of comparative advantage constrained its space to focus on spatial growth irrespective of the boundaries of nations and their countries. By implication, the basic assumptions of traditional trade theory – such as the immobility of productive factors across the countries, their aggregation into homogeneous categories such as land, labour and capital, the presence of constant returns to scale, the absence of transport costs, and the use of non-tradable goods as inputs, all have created hardly any space to approach the issues of spatial growth concentration. Interestingly, as we have already discussed the authors of the Heckscher-Ohlin trade theory have spent substantial amount of space in their initial work to observe and analyze trade implications of all these assumptions which they did not take as given conditions of trade theory.

This is, however, not to assert that the explicit or implicit assumptions of trade theory never became the topics of interests in subsequent developments of the discipline. On the contrary, they were part of the academic discussions and debates, and guided even the theoretical extensions. While the presence of increasing returns to scale and its implications on market structures and international trade which received considerable attention in trade analyses since the 1960s, became the foundation of new trade theory (Krugman 1979). The transport cost and physical distance entered into trade analyses as a major determinant of bilateral trade following the development of Gravity model (Tinbergen 1962). In spite of that, to a great extent the assumptions that were meant to simplify the complex matters at the beginning continued also to be conditional in theoretical and empirical studies. Apart from that, the intellectual drift away from ‘modeling the world’ to ‘developing the models’ that came to dominate mainstream economics led to ignore even the important analytical issues as such from mainstream economic theory (Bandara and Jayasuriya 2011). As a result, after the early contributions of Heckscher and Ohlin to trade theory and to the analysis of the implications of its underlying assumptions, the nexus between trade and spatial growth appear to be missing in trade analyses.

4 Development Perspectives on Spatial Growth

What was missing in trade analyses was, however, not missed by the economists who worked within the broader context of economic development from different perspectives. There has been concern in early development literature about the growing economic inequality across geographical space arising from national income growth as well as international trade. Some of the early contributions to the study of
spatial concentration of production and specialization have come from urban economics, regional economics, and the analyses of industrial location. Even though ‘location of production’ was not treated as an equally important issue in mainstream trade theory, it did not remain entirely out of sight in the field of economics.

The spatial concentration of growth and the resulting regional inequality in the development process received much attention of early economists such as Alfred Marshall, Joseph Schumpeter, Allyn Young, Albert Hirschman, Gunar Myrdal, Nicholas Kaldor, and François Perroux. To a certain extent, the topic was also of the interest of development theorists in the dependency and neo-Marxist traditions which were based on the notion of ‘center-periphery’ exploitative relationships.

It was Myrdal who conceptualized ‘circular and cumulative causation’ which can operate in either direction leading to a rise of economic concentration in some locations and to a fall of that in other locations. Myrdal (1957:17) explained that the whole system that starts moving in one direction due to changes in the forces working in the same direction, because the variables are so interlocked in a circular causation mechanism; ‘...a change in anyone induced the others to change in such a way that these secondary changes support the first change, with similar tertiary effects upon the variable first affected, and so on.’ He acknowledged that migration, capital movements and international trade are rather the media through which the circular and cumulative process evolves.

The circular and cumulative process in the growing regions which are the lucky regions would thwart the other regions – the unlucky ones, which enter the downward path of the circular and cumulative causation:

If things were left to market forces unhampered by any policy interferences, industrial production, commerce, banking, insurance, shipping and, indeed, almost all those economic activities which in a developing economy tend to give a bigger than average return – and, in addition, science, art, literature, education and higher culture generally – would cluster in certain localities and regions, leaving the rest of the country more or less in a backwater. (Myrdal 1957:26)

This is true enough, and reflects the outcome of circular and cumulative causation. However, it means that Myrdal did not perceive that spatial growth concentration could have a positive impact on lagging regions through connectivity or that regional disparities could converge at latter stage of development after they diverged at its early stage.

Nicholas Kaldor, referring to Myrdal’s concept of circular and cumulative causation, attributed it nothing more than to the presence of increasing returns to scale, both internal and external to the firm:

This is nothing else, but the existence of increasing returns to scale – using that term in the broadest sense – in processing activities. These are not just the economies of large-scale production, commonly considered, but the cumulative advantages accruing from the growth of industry itself – the development of skill and knowhow; the opportunities for easy
communication of ideas and experience; the opportunity of ever-increasing differentiation of processes and of specialization in human activities. (Kaldor 1970: 480)

While economies of large scale production, as referred to by Kaldor, means the presence of increasing returns to scale due to the expansion of the productive capacity of a firm as permitted by the larger size of input and output markets. In addition, the term returns to scale is used in its broadest sense to capture the advantages arising from the externalities within the business environment in which the firm is located.

According to Hirschman (1958), growth is necessarily unbalanced, and it must be so in order to lift a country to a higher income levels because it has to first develop within itself one or several regional centers of economic strength. The ‘...need for the emergence of “growing points” or “growth poles” in the course of the development process means that international and interregional inequality of growth is an inevitable concomitant and condition of growth itself’ (Hirschman 1958:183-184). Growth poles within a country emerged and progressed through backward and forward linkages of industry so that input-output relations of industry can guide the formulation of development strategy.

A distinguished and influential branch of modern theoretical developments on spatial growth was centered on François Perroux’s ‘growth pole theory’, spurring debates and discussions on the concept, theory and its policy implications in the 1960s and the 1970s. Perroux viewed that economic space of a firm is more important than its banal space, and that this economic space is a field of centrifugal and centripetal forces:

As a field of forces, economic space consists of centres (or poles or foci) from which centrifugal forces emanate and to which centripetal forces are attracted. Each centre being a centre of attraction and repulsion, has its proper field, which is set in the fields of other centres. Any banal space whatever, in this respect, is a collection of centres and a place of passage for forces. (Perroux 1950:95)

Being a critique of mainstream economics, François Perroux believed in the role of the state in economic development. Thus, his growth pole theory guided the contemporary regional development planning and state-led resource allocation programmes. As early as late-1960s Perroux observed that ‘growth does not appear everywhere and all at once; it reveals itself in certain points or poles, with different degrees of intensity; it speeds through diverse channels’ (Higgins and Savoie 1988: 6). It is necessary to identify such growth poles in regional development planning, and fulfill the conditions to accelerate self-sustained economic growth process. The approach has guided regional development policies and planning exercises in most of the advance countries during the early post-war period.

A wide range of dissent views on growth disparities across the space could be found in contemporary trade and development literature emerged in the neo-Marxist and dependency traditions particularly during the early postwar period. The work of Raul Prebisch and Hans Singer contributed initially to the development of the dependency theory, which shared much in common with the neo-Marxist economists such as Paul Baran, Paul Sweezy, Samir Amin, and Arghiri Emmanuel.
The theories were primarily presented in international context, focusing on the exploitative or unequal trade relations between industrialized and underdeveloped countries. The industrialized countries were playing the role of the ‘centre’ and the underdeveloped countries that of the ‘periphery’. Even within a country, the metropolitan areas that are connected to the ‘centre’ play a mediating role accumulating resources and surplus from the remote peripheral areas of the country resulting in a regional polarization within the countries. Regardless of geographical scale, the ‘centre-periphery’ relations through capital accumulation in the centre and unequal exchange between the centre and the periphery lead to a polarization of regions across the space. By implication, therefore, studies in this tradition established rationale for delinking the periphery from the centre in order to initiate a self-sustaining and independent growth process. In addition to their influence in the contemporary global political spheres, these economic perspectives had a bearing impact on shaping import-substitution and state interventionist strategies in developing countries during the early post-war period.

The academic vigour and policy relevance of the perspectives as such did not continue to remain as strong as they used to be during the early postwar period. They were weakened, on the one hand by the revival of neoclassical Economics and trade liberalization reforms among developing countries since the 1980s. On the other hand these developments in the field of International Economics were supported by the contemporary empirical evidence on industrial transformation and growth concentration in newly industrialized countries in East Asia against dismal growth performance in most of the developing countries.

5 Trade and Spatial Growth, Revisited

Global economic dynamics created by liberalization reforms and global integration since the 1980s opened up the space within international economics which was once closed by modeling assumptions, to look into intra-national economics of spatial growth. This period has been marked by dramatic changes in the world economy, which have exhibited implications on the spatial concentration of growth at every form of geographical scales – global, regional, and local. The reforms of trade liberalization and market deregulation have led to a greater integration of countries with the global economy leading to trade expansion, financial flows and capital movements. This was further reinforced by the formation of regional blocs within which integration began to speed up at varying degrees from free trade areas to economic and monetary unions.

An important turning point in regional integration among advanced countries was the progress of the European Monetary Union, and the establishment of the Euro Zone. The free movement of goods and services, as well as capital and labor, led to an enhancement of trade in goods and factors and a reduction in transport cost and distance within the Euro Zone. More importantly, the withdrawal of cross-border barriers of factor mobility called for revisiting the trade theories of comparative advantage in a new global environment. As the World Bank (2009) observed, thus greater integration in the world has shortened the distance, and reduced the divisions, facilitating the countries and regions to accelerate and concentrate their economic growth.
World merchandise exports shared by advanced and developing countries have changed with rapid export growth over the past few decades. Merchandise exports from developing countries have increased 15 times during the period of 1985-2012, compared to those of advanced countries by 7 times. The share of exports from developing countries was around 30 percent of world total during the 1980s (Figure 2.1). This has gradually and steadily increased so that by 2012 the advanced and developing countries account for equal shares of world exports. Among the developing countries, the exports from the Asian region accounted for the highest increase in its share.

An important feature of exports expansion from the developing region in general, and from Asian region in particular has been the change in trade patterns deviating from ‘old-styled’ international trade. It is no longer the type of exchange between ‘wine and cloth’ as in traditional inter-industry trade models of comparative advantage, and neither the exchange of ‘wine for wine’ or ‘cloth for cloth’ as in intra-industry trade models. A new facet of growing trade patterns is the global product sharing, led by the production and export of ‘parts and components’ by different countries, followed by assembling activity and export of the final good by another country.

The countries that have been integrating into global supply chains have been specializing in ‘performing tasks’ instead of ‘producing goods’ (WTO 2011). The process has led individual countries to gain from trade in tasks according to a different form of comparative advantage. Consequently, the rapid growth of ‘network trade’ and the rising shares of parts and components in total exports have been an outstanding feature of the export-oriented growth process in East and Southeast Asia. While trade in parts and components and assembled products has generally grown faster than total world trade in manufacturing, East Asia’s dependence on this new form of trade pattern is proportionately larger than elsewhere in the world (Athukorala 2011). Although network trade has been growing in South Asia too, even with regional and bilateral trade agreements, countries such as India and Sri Lanka have performed marginally in respect of product sharing within the region (Abeyratne 2013).

The slicing up of the value chain of a commodity into finer parts and components, and international outsourcing of their production and supply have been made possible by a series of factors that came into operation at the same time simultaneously. It is the TNCs which began to relocate their investment and production from advanced countries to developing countries that played the major role in the fragmentation of the production processes and the formation of globalized value chains. The globalized value chains are typically coordinated by TNCs with their cross-border networks of affiliates, contractual partners and arm’s-length suppliers, while the TNC-coordinated global supply chains today account for some 80 per cent of global trade (UNCTAD 2013:xxii). Secondly, the technological developments, which were undertaken by the TNCs, opened up the technical feasibility of network trading. These technological developments in two basic dimensions allowed slicing up the value chains into finer parts and components on the one hand, and facilitated their international outsourcing due to a dramatic decline in transportation and communication costs on the other hand. Finally, the greater integration of different countries and regions through policy reforms towards liberalization and deregulation reduced the cross-border barriers to network trade.
Historically the world FDI flows were used to be more within the region of advanced countries than from advanced to developing countries. The past few decades were marked by a dramatic change in FDI flows in the world. While there has been a phenomenal growth of world FDI flow and its stock, developing countries began to account for an increasing share of it. As UNCTAD (2013:217) reported, the world FDI stock has increased from USD 2 trillion in 1990, to USD 7.5 trillion in 2000, and to USD 22.8 trillion in 2012. The share of world FDI flows into developing countries has increased and outnumbered that of the advanced countries (Figure 2.2). In the 1980s and the 1990s, the FDI flows to high income OECD countries were generally in the range of 60-80 percent of the world total, which has declined to below 50 percent during the first few years of the 2010s. Correspondingly, the share of FDI flows to developing countries increased, and accounted for a record 52 per cent of global FDI inflows, exceeding flows to developed economies for the first time ever, by $142 billion (UNCTAD 2013: xii). The rising magnitudes of the world FDI flows, and their flowing directions have confirmed that the Transnational Corporations (TNCs) have been relocating their investment and production from advanced countries to developing countries in general, and to developing Asia in particular.

The exponential growth of the world FDI flows and their increasing diversion towards developing countries were contributed by a number of contemporary economic phenomena. It was the time that developing countries started to adopt policy reforms and offered investment environs conducive to FDI inflows. Coincidently, the advanced countries had ended their postwar growth momentum, entering into a long phase of economic slowdown in which the FDI outflows started to rise more than ever before. Simultaneously, the advanced countries also adopted policy reforms aimed mostly at liberalizing the financial and foreign exchange markets which facilitated international financial flows through the removal of barriers to trade and investment. The technological advancement which was originating from the TNCs led a fragmentation of their production processes and their relocation across many countries resulting in new trade patterns. Thus location of production emerged as one of the central issues in trade analyses.

The use of trade theory in analyzing the location of production and specialization made a new turn in the early 1990s with the publication of Krugman (1991), which claimed to be the beginning of new economic geography: ‘It is almost exactly twenty years since I delivered a set of lectures in Leuven in Belgium that became the monograph Geography and Trade, which most people consider the beginning of the New Economic Geography.’ (Krugman 2011:1). However, there have been a number of theoretical contributions and policy analyses in the field of economic geography particularly since the early 2000s (Fujita et. al. 2001, Fujita and Krugman 2004, Venables 2009, World Bank 2009), apart from country-specific empirical analyses. The goal of new economic geography was to devise a modeling approach to location of production, by incorporating and amalgamating its key elements, as following:

The first is the general-equilibrium modeling of an entire spatial economy, which sets apart our approach from that of the traditional location theory and economic geography. The second is increasing returns or indivisibilities at the level of the individual producer or plant, which is essential for the economy not to degenerate into “backyard capitalism” (in which each
household or small group produces most items for itself). Increasing returns, in turn, lead to the market structure characterised by imperfect competition. The third is, of course, transport costs (broadly defined), which makes location matter. Finally, the locational movement of productive factors and consumers is a prerequisite for agglomeration. (Fujita and Krugman 2004: 142)

The key aspects of new economic geography approach to location of production have been in the heart of trade and development theories. According to Ohlin’s (1933) detailed analysis, the conditions of factor mobility, cost of transportation of both productive factors and commodities, and the cost advantage of large-scale production exhibit underlying conditions of location theory – concentration of production and specialization. In fact the contributions to new trade theory since the late 1970s had already incorporated economies of scale into trade modeling, which showed that ‘...trade need not be a result of international differences in technology or factor endowments’ (Krugman 1979: 479). Instead, trade grows with the expansion of the market and the exploitation of economies of scale within imperfectly competitive market structures. Although in new trade models it was the internal economies of scale that lead to imperfect market structures and product differentiation, as Ohlin (1933) recognized both internal and external economies of scale play an equally important role in the location of production and specialization.

There are both centripetal and centrifugal forces operating in specific locations, as conceptualized in Perroux’s (1950) growth pole theory, and as became the cornerstones of the modeling of new economic geography. Centripetal forces attract economic activities together, when the increasing returns to scale are present and when the transportation has a cost. When both factors are taken together, a decline in transport costs enhances spatial growth concentration and expands trade across locations of production. This is what has happen with technological advancement leading to a reduction in transportation costs and product fragmentation, together with trade liberalization and global integration.

Then firms have an advantage of locating themselves closer to their suppliers (input markets) and buyers (output markets). Hirschman’s (1958) backward and forward linkages, and Myrdal’s (1957) circular and cumulative causation which was reinterpreted by Kaldor (1970) as returns to scale, all explain why centripetal forces bring together economic activities and bring about spatial growth concentration into a self-sustaining stage.

In contrast, centrifugal forces disperse economic activities away from each other, when the costs of agglomeration outnumber its advantages and when the productive factors are immobile. Agglomeration also has costs in terms of higher factor prices resulting from rising demand, inelastic supply, and the costs of non-tradable goods, as well as the costs of congestion. As Ohlin (1933) also observed, economies of large scale production could lower part of these costs. However, the types of industries matter too, as some of the economic activities (such as those in agriculture, resource-based industries, and some tourism-based industries) do not concentrate in the same locations as many others do. While natural resources categorized under land are an immobile factor, as Ohlin (1933) explained, industry demand is mostly for a set of factors which is less mobile. Depending on the type of industry demand for a set of factors could be seen as operating in both directions forming centrifugal as well as centripetal forces. Therefore, it is a reason for some economic activities to disperse, and for others to concentrate.
World Bank’s (2009) ‘three-dimensional’ approach to geographical concentration of economic activity is based on the concepts of density, distance, and division. Out of the three dimensions defined in economic terms, the first stands to be the outcome of the latter two which of course appear as means, and not ends. Economic density, which is defined and quantified as the economic mass per unit of land area or the geographic compactness of economic activity (World Bank 49) shows how growth has concentrated or dispersed over geographical space. The other two concepts – distance and division, are some of the important drivers of economic density. The economic distance is ‘...the ease or the difficulty for goods, services, labour, capital, information and ideas to traverse space’ (World Bank 2009: 75). Thus it is a broader definition of distance, which is not limited to physical distance only. It includes all dimensions of distance over space affecting differently for different things to traverse. The third concept of division in economic sense, i.e. economic division, includes a range of restrictions on the flow of goods, capital, people, and ideas as well as more severe divisions caused by political disputes and security issues (World Bank 2009: 97). Therefore, reforms towards trade liberalization and deregulation are an important step in the direction of reducing economic divisions between and within countries.

6 Centripetal and Centrifugal Forces of Spatial Growth

Whatever the ‘geographic scale’ of the unit of analysis – international or intra-national, in response to the process of policy reforms growth was seen as accelerating and concentrating. While integration of nations, as we have already discussed, has called for trade analyses in different forms, it has also paved the way for looking at the location of production and specialization within as well as without the territorial boundaries of countries. In the case of trade blocs and agreements too, a greater degree of policy coordination has emerged as a requirement of the integration of nations which has been progressing far deeper into economic and monetary unions. Factor mobility is largely free within a country, although divisions and restrictions could still exist mostly beyond economic spheres. In the progressive stages of integration too, national borders that restrict factor mobility, including labour mobility, have begun to disappear.

Even though in trade models the assumptions were relaxed and specifications were changed, the fascinating outcome of trade liberalization and integration has been that it still accelerates trade, and hence promotes growth, as evidenced in high-performing economies in many parts of the world. Secondly, growth concentrates, and that concentration tends to be faster in high-performing economies than in other countries. Policy reforms and technological progress that have expedited global integration lead to a decline in costs of trading across nations, but make centripetal forces stronger than centrifugal forces; in other words, when trade is growing faster growth is concentrating more.

The theoretical underpinnings of spatial growth based on international economics since the time of the birth of trade theories could be classified under policy-oriented four broader headings: benefits of agglomeration, costs of connectivity, factor mobility, and market size.

As the benefits of agglomeration, the economic advantages of ‘being together’ that lead to a concentration of economic activity and people in particular locations. These benefits can be approached
in two respects as those accruing to the individual economic activities or households at micro level, and as those accruing to the country or nation at macro level. The presence of internal and external economies of scale is an important mechanism of receiving the benefits of agglomeration by firms and industries. The developmental achievements through trade expansion and rapid economic growth as well as improvements in related components are the benefits aggregated at macro levels. People also agglomerate as suppliers of labour and consumers of output, as well as the beneficiaries of agglomeration. The choice of policies can operate in either direction, in facilitating or hindering the benefits of agglomeration to firms and households and thereby to the national economy.

Connectivity of a particular location is an important part of the cost of economic activity concentrated in that location. The costs of connectivity are associated with connecting the economic activity with various types of input and output markets, and are measured in terms of monetary costs, time delays, and quality of traversing. A wide range of connectivity costs, as outlined in World Bank (2009), explains how efficient and effective is the connectivity of an economic activity to various markets such as labour markets, financial services, business services, regulatory mechanisms, information sources, communication networks, energy sources, and transport corridors. Transport cost is, in fact, an important cost component of connectivity, but the overall connectivity grid of an economic activity is not limited to physical distance or transport costs. Some of the important connectivity costs need to be identified as emanating from a wide range of non-tradable service sectors such as trade and commerce, seaport and aviation services, domestic transport, information and communication, personal and social services, most of the health and education services, utility supply, law and order, security, and administrative and regulatory services. One way of reducing the costs of connectivity is, in fact, agglomeration – a choice made by the firm, but not limited to the agglomeration alone because policy environment and infrastructure also account for, perhaps a greater share of it.

It is clear that factor mobility is not perfect internally or internationally, but from an economic point of view the benefits of agglomeration and the costs of connectivity account for its greater flexibility. Given the broader classification of all productive factors into homogeneous categories, capital is almost freely mobile, labour is more flexible internally than internationally, and land is immobile. The world experience suggests that global capital which freely traverses across the world concentrates in some locations of some countries only. Where capital accumulates, human resources concentrate too. The policy choices can either accelerate or decelerate the capital and labour concentration. Education and human resource development can encourage spatial concentration of labour, thereby facilitating growth as well as specialization of production. On the contrary, the policy choices can discourage spatial concentration of labour as well as capital, but at a cost to the firms and households as well as to the economy as a whole. At the same time it is necessary to acknowledge the immobility of some factors as well as that of a ‘set’ of factors, but they could also generate centripetal forces of spatial growth. As industry demand is often for a set of different factors in terms of volumes and types; for instance, different categories of labour skills and their adequate quantity do not move together so that a set of factors as such is not be available evenly across the geographical space. Apparently, natural resources and locational advantages constitute the sources of regional comparative advantages within a country, and create spatial growth impetus.
The size of the market limits the extent of concentration of economic activity and specialization of production as well as the capacity for deriving benefits from economies of scale. Trade liberalization expands the boundaries of the market, resulting in both an acceleration of growth and its spatial concentration. Centripetal and centrifugal forces of spatial growth become stronger in the context of liberalization policy reforms as firms shift their focus from limited domestic markets to international markets. Therefore, access to larger international market though ‘international gateways’ is an important factor underlying spatial concentration of economic activity. For this reason, most of the outstanding economic agglomerations in the world are located in and around international gateways to global markets that have transformed the respective locations into global hubs of economic activity. The point is more important for smaller countries which have to worry more about the size of their domestic markets than the larger ones which can rely on their larger domestic market at least for a relatively longer period.

When the issue of regional growth concentration and specialization is addressed in the context of trade theory, it is not necessarily about the specialization and trade between two regions of an isolated country in a closed economy model. It is an open economy within which liberalized trade regime leads to an acceleration of international trade and economic growth, and it simultaneously shapes spatial location production and specialization. Unevenness in spatial growth is an inevitable outcome of trade-growth nexus, due to higher benefits of agglomeration, lower cost of connectivity, greater degree of factor mobility, and the larger size of the market. The choice of policies and reforms can work in either direction in facilitating or constraining spatial growth concentration. Particularly, the policies need to contribute to the centripetal forces, and not to the centrifugal forces which could disperse economic activities and people at a cost to them as well as to the economy.

7 Concluding Remarks

The theories of comparative advantage are based on the location-specific factors as the determinants of trade. This was not strange to Bertil Ohlin who considered trade theory as part of location theory, and to his teacher who might have inspired him, Eli Heckscher – the co-authors of Heckscher-Ohlin trade theory; they did not get locked up in the assumptions of trade theory. What prevented trade analyses capturing spatial growth were the strict assumptions of traditional trade theory of comparative advantage. As a result, for decades the trade theory and location theory developed independent of each other as two separate disciplinary branches. Although trade was less explicit than it was in Ohlin’s analysis, location of production was not a peculiar issue to many of the early economists from divergent traditions. Mainly within the premises of development economics, they conceptualized the location of production by focusing on the factors which occupied the central position of economic geography.

As the world began to change in its shape due to greater integration and globalization within which trade expanded and growth accelerated, most of the assumptions of trade analysis became obsolete and space for location analysis within international economics resurfaced. The economic analysis within international economics is a not only fills the gaps in knowledge, but also contributes to emerging policy debates on trade and spatial growth. The main factors underlying the spatial growth concentration can
be categorized under the benefits of agglomeration for firms and people, costs of connectivity to input and output markets, the degree of mobility of the non-homogeneous productive factors, and the size of the markets. Given the nexus between trade and spatial growth, it is the choice of trade strategy that appears to be far more important than anything else in accelerating and in reshaping spatial growth within a country. This was not missing in the initial Heckscher-Ohlin trade theory, although it was missed by many (but not all) of the subsequent trade analyses.
References


WTO (2011), Trade Patterns and Global Value Chains in East Asia: From Trade in Goods to Trade in Tasks, Geneva: World Trade Organization
Figure 2.1 Shares of World Exports in Developed and Developing Countries, 1980-2012

Note: Classification of developed and developing countries is according to the data source, while developing countries include transition economies.

Figure 2.2 Shares of World FDI flows to Developed and Developing Countries 1980-2012

Note: Classification of developed and developing countries are according to the data source, while developing countries include transition economies.