# Developments of New Economic Geography: From Symmetry to Asymmetry*

Jian WANG
Xiao-Ping ZHENG

**Abstract**

This paper presents an overview on the recent developments of the so-called New Economic Geography (NEG), which successfully applies the basic microeconomic principle to explain the endogenous formation of spatial agglomeration in the field of urban and regional economics. The paper provides a detailed explanation of Krugman's Core-Periphery model (1991) and reviews the subsequent refinements of this original model mainly from geographical viewpoints. It points out that most of the existing NEG models are symmetric in that regions or countries are assumed to have the same sizes and accessibilities to markets, which is not realistic and is unable to explain the real spatial economy. The paper concludes that it is important for NEG to take into account the full asymmetry in terms of different regional sizes and accessibilities so that it can adapt its theories to more accurately reflect the real world.

**Key words**: New Economic Geography, symmetry and asymmetry, spatial agglomeration, trade costs, transport costs

## 1. Introduction

Economic activities are always relevant to the two dimensions of time and space. While time is often integrated into economic theoretical considerations, the study of where economic activities take place is almost always ignored by mainstream economists. In fact, in the real world, spatial distributions of population and employment are never smooth. In addition to the uneven endowment of natural resources, economic forces called the second nature also play a dominant role in shaping economic geography. Though this subject was once regarded intractable, some exciting new models dealing with increasing returns and monopoly competition have emerged (Dixit and Stiglitz, 1977). In the wave of increasing-returns revolution, the new trade and new growth theory are born successively. Based on

* The authors wish to thank the Japan Society for the Promotion of Science (JSPS) for providing Grants-in-Aid for their research projects (No. 23530300).

them, Paul Krugman's seminal paper "Increasing returns and economic geography" (1991) endogenously reaped the mechanism of economic agglomeration and inaugurated the new era of spatial economics, which became known as New Economic Geography (NEG). Since then, a large amount of improved works under this framework have been developed to analyze the impacts of globalization and regional integration on the spatial economy.

It has been two decades since the birth of New Economical Geography. Its theoretical merits have encouraged mainstream economists to pay much more attention to the spatial dimension of the economy. NEG has been successful in offering rationales for the formation and existence of economic agglomeration or the cluster in geographical space. NEG provides an equilibrium framework to explain the market access forces pulling economic activities together and the market-crowding ones pushing them apart, resulting from the tradeoffs between increasing returns and mobility costs. Compared to its antecedents such as regional science and urban economics, NEG attempts to enter the black box of agglomeration economies and show the self-reinforcing character of spatial concentration from more fundamental considerations (Fujita et al. 1999, p. 4). Another merit of NEG is that the formation of economic agglomeration is derived explicitly based on a full micro foundation. Economic space is represented as the outcome of the equilibrium individual's interaction.

NEG's achievements attract more and more attention from mainstream economics. However, despite its countless refinements and rapid progress following Krugman's seminal paper, NEG still has some internal and external drawbacks and challenges. During the glorious process of advance, queries have always accompanied the progress. One question for NEG is that many of its stories about industrial agglomeration or cluster cannot explain the real countries and regions. The mismatch between its original theory and the real space-economy, as Krugman (2011) puts it, means that the NEG model has entered middle age and seems decreasingly applicable to actual location patterns of advanced nations. More than that, economic geographers complain that NEG overemphasizes the role of pure economic mechanism and neglects the complexities of geographical features. Even though the model is simple and illuminating, such models are unable to explain the rich and complex hierarchy that characterizes the space-economy (Ottaviano and Thisse, 2004). NEG models really tell the story about why economic agglomeration occurs, but they have little concern about where the agglomerations arise. Most NEG models assume that there are only two regions and the distributions of economic activities are decided by pure economic power. In the real world, however, regional comparative advantages still play a non-trivial role through their interaction with the second nature (Fujita and Mori, 2005). In NEG, homogeneity of location space is facilitated to investigate the complicated mechanism of economic agglomeration. However, it is too unrealistic to explain the real uneven world. The landscape of economy forms under a synthesis of the first nature and the second nature.

Nowadays, globalization has brought about manufacturing outsourcing and spreading from developed countries to developing ones. The emerging economies, especially the

BRICs (Brazil, Russia, India, and China) countries, shoulder the main task of manufacturing production for the world. These economic powerhouses play an important role in the world's economic stage. When investigating the space-economy within these countries, economic geography seems reminiscent of the circumstance of developed nations at the beginning of the twentieth century (Krugman, 2011). Economic landscape in these nations is reshaped in the process of international trade liberalization. For example, China's industrial growth and agglomeration has been accompanied by massive migrant labor from its hinterland to coastal regions where manufacturing industries are massively concentrated. Empirical evidence reveals that the regional disparity is increasing due to the economic liberalization and globalization between 1985 and 1994 in China (Fujita and Hu, 2001). This means that the concept and approach of NEG is not outdated after all. China's industrial cluster and localization in its eastern region are the evidence to support NEG.

In addition, regional developments under economic integration exhibit differentiated responses across regions within a country. With the development of economy, especially the economic integration and the improvement of infrastructure within these countries, the relocation of manufacturing to hinterland is increasing in developing countries such as Mexico and Brazil (Hanson, 1996; Henderson et al., 2001). In China, the growth rate of per capita GDP exhibited some convergence after the large gap from the late 1990s to 2006 between hinterland and coastal regions (Fan and Sun, 2008). In addition to local government policy, industrial dispersion from east to west is a growing tendency in China. Fally et al. (2010) find that market and supplier access have a stronger positive impact on wage disparity across Brazil, and that there is higher population density in Brazil's coastal regions. Similarly, in a study of post-Soviet Leningrad Oblast, Golubchikov finds that in the integration under globalization, the geographical elements of region proximity and size is responsible for the formation of economic core and periphery (Golubchikov, 2006). In these cases, the original core-periphery theory appears to be weak to explain the phenomenon of industry redispersion.

It is important to note that what happened in recent years in the real world proposes some new issues for NEG. The distribution of economic activity within a country is affected by the integration of international trade liberalization and the boost of domestic infrastructure simultaneously. Moreover, the domestic regions in these countries are asymmetrical in terms of their access to the overseas market. There are well-documented facts that the market accessibility of domestic regions has a significant impact on its trade volume and firm agglomeration. More importantly, the hinterlands in emerging economies are always endowed with more unskilled workers. Although NEG sets forth a pure economic theory to explain the formation of agglomeration, the weakest point of it is its geography. It is time to consider this limit in the explanation of spatial economy. As noted, earlier NEG literature focused on self-reinforcing mechanisms in reshaping spatial agglomeration. It is time for us to incorporate the omitted asymmetrical geographical features into the original theory to explain the increasing regional disparity in emerging countries in the process of globalization. It seems necessary to step out the assumption of only two regions,

two industries and the iceberg transport cost.

Based on the above considerations, the purpose of this paper is to present a detailed review of the refinements in NEG. As there are several profound surveys on the developments of NEG, we focus especially on geographic space. Early NEG models consider only two regions or countries that are symmetrical in terms of size and market access. After 20 years of developments, there appears an evolution from symmetry to asymmetry in NEG. The models having asymmetric assumptions begin to emerge. These models are becoming increasingly sophisticated. It is essential to sort out their typical papers and make some summarization. Although some works have attracted sufficient attention, we need to shed light on their future possibility. The remainder of the paper is organized as follows. We first describe the basic framework of NEG models in Section 2. To show the related developments of NEG, we review the existing works and compare their differences in Section 3. In Section 4, we provide a further research direction to consider full asymmetry. We emphasize that the first nature advantage of regions plays a noteworthy role through its interaction with pure economic forces. If domestic regions are endowed with different regional scale and access to the world markets, a theory based on these assumptions will be much closer to reality. Finally, concluding remarks are given in Section 5.

## 2. The Framework of New Economic Geography

Krugman's seminal Core-Periphery (CP) model (1991) provides a basic framework for most NEG models. It clarifies how interactions among increasing returns and transport cost can result in the emergence and evolution of economic agglomeration. In this section, we briefly review the basic NEG models. Although there are abundant extensions to the CP model, the structure of them is identical (Robert-Nicoud, 2005). We also introduce those extensions that step outside the canonical framework. Employing a quasi-linear utility function and linear transport costs, Ottaviano et al. (2002) bring fresh air to NEG, presenting analytical results regarded as a great breakthrough in NEG.

### 2.1 The Core-Periphery model

The classic CP model (Krugman, 1991) is considered the basis of NEG models. In it, the spatial economy consists of two regions that are evenly endowed with the same economic structure. There are two sectors in each region: one is the numéraire (e.g., agriculture) and another is manufacturing, respectively denoted $A$ and $M$. The numéraire sector produces a unit of homogenous goods with constant-return-to-scale technology exploiting one unit of unskilled labor under perfect competition. A representative firm of the manufacturing sector supplies differentiated goods using skilled workers under monopolistic competition. All goods and skilled workers are mobile between the two regions, but unskilled workers are immobile.

(1) **Consumption**

A representative consumer in each region has the same CES (constant elasticity of substitution) utility function in the following form:

$$U=C_M^{\mu}C_A^{1-\mu},\quad C_M=\left[\int_0^n q(i)^{1-\frac{1}{\sigma}}di\right]^{\frac{\sigma}{\sigma-1}}$$

where $q(i)$ and $C_A$ denote the consumption of manufacturing goods variety $i$ and numéraire goods, respectively, and $\mu(0<\mu<1)$ is a constant denoting the expenditure share of manufactured goods. $n$ is the range of varieties produced, and parameter $\sigma(\sigma>1)$ is the constant elasticity of substitution between any two varieties. $1-1/\sigma$ represents the intensity of the preference for variety in manufactured goods.

The consumer has a budget constraint given by $p_AC_A+\int_0^n p(i)q(i)di=Y$ where $p_A$ and $p(i)$ are prices of numéraire goods and manufacturing goods variety $i$ respectively; $Y$ is his income. The consumer's problem is to maximize the utility function subject to the budget constraint by choosing his amount of consumption.

Following Fujita et al. (1999), we express the consumer's demands for numéraire and manufacturing goods as follows:

$$C_A=\frac{Y}{p_A}(1-\mu) \tag{1}$$

$$q(i)=\mu Y\frac{p(j)^{-\sigma}}{P^{-(\sigma-1)}} \tag{2}$$

where $P$ is the price index for manufactured goods having the following form:

$$P=\left[\int_0^n p(i)^{1-\sigma}di\right]^{1/(1-\sigma)} \tag{3}$$

As is assumed in the standard models of New Economic Geography, each variety of manufacturing goods is produced in only one region where firms produce them with same technology and price. The manufacturing goods are traded between the two regions, which is associated with transport costs. The numéraire goods are freely transported between the regions, and their prices are equalized.

The iceberg form of transport costs is assumed for manufacturing goods, meaning that when goods are shipped from region $r$ to region $s$, it is assumed that only a fraction $1/\tau_{rs}$ arrives. Here, $\tau_{rs}$ represents such a unit transport cost. It means that the price of goods consumed in region $s(p_{rs})$, which are produced in region $r$ at price $p_r$, can be written as $p_{rs}=p_r\cdot\tau_{rs}$. We denote the number of varieties produced in region $r(s)$ by $n_r(n_s)$. Using equation (2), the demand function of consumers in region $s$ for manufacturing goods produced in $r$ can be expressed:

$$q_{rs}(j)=\frac{{p_{rs}}^{-\sigma}}{P_s^{1-\sigma}}\mu Y_s \tag{4}$$

Then the price index in region $s$ can be written as:

$$P_s=[n_r(p_r\cdot\tau_{rs})^{1-\sigma}+n_s p_s{}^{1-\sigma}]^{\frac{1}{1-\sigma}} \tag{5}$$

The total demand for the variety produced in location $r$ can be expressed as follows:

$$q_r=\mu\frac{p_{rs}{}^{-\sigma}}{P_s{}^{1-\sigma}}Y_s\cdot\tau_{rs}+\mu\frac{p_{rr}{}^{-\sigma}}{P_r{}^{1-\sigma}}Y_r \tag{6}$$

### (2) Production

We now turn to the production side of the economy. The numéraire goods are produced with constant-returns-to-scale technology under perfect competition. We assume that one of unskilled workers produce only one unit of numéraire goods, and workers' wages in the two regions are equalized at one. To produce $q(i)$ amounts of manufacturing goods at region $r$, $l^M$ amounts of labor are used, i.e.

$$l_r^M=\alpha+\beta q_r(i) \tag{7}$$

where $\alpha$ and $\beta$ represent the fixed and marginal costs of manufacturing, respectively. The firm in region $r$ determines its outputs so as to maximize its profit, written as follows:

$$\pi_r=p_r(i)q_r(i)-w_r(\alpha+\beta q_r(i)) \tag{8}$$

where $w_r$ is the wage rate. Given the previous assumptions, profit maximization of the firm results in the price of its output as follows, and the price in region $s$ can be written as

$$p_r=\frac{\sigma}{1-\sigma}\beta w_r \tag{9}$$

$$p_{rs}=\frac{\sigma}{1-\sigma}\beta w_r\cdot\tau_{rs} \tag{10}$$

Due to the free entry into market, the zero-profit condition exists, which yields the equilibrium output of the firm as follows:

$$q_r^*=\frac{\alpha(\sigma-1)}{\beta} \tag{11}$$

And the input of labor can be written as

$$l_r^M=\alpha+\beta q_r=\alpha\sigma \tag{12}$$

Denoting $L_r^M$ to be the total number of manufacturing workers and the number of firms in region $r$ by $n_r$,[1] it can be written as:

$$n_r=\frac{L_r^M}{l_r^M}=\frac{L_r^M}{\alpha\sigma} \tag{13}$$

Concerning the meanings of (4) and (13), some additional interpretations are worth putting forward. The number of varieties produced in a region is proportionate to the regional

whole labor force. If we assume that all manufacturing goods are available at the same price, then from equation (4), $P$ can be simply expressed as $P_r = p_r^M n^{1/1-\sigma}$. It means that the increase of manufacturing varieties will reduce the price index.

(3) **Short-run equilibrium**

For the manufacturing goods, using (6), the supply-demand balance condition gives:

$$q_r^* = \mu \frac{p_{rs}^{M-\sigma}}{P_s^{1-\sigma}} Y_s \cdot \tau_{rs} + \mu \frac{p_{rr}^{-\sigma}}{P_r^{1-\sigma}} Y_r \tag{14}$$

By using the equilibrium price $p_r$ given by (9) and the NEG's traditional normalization of $\beta = (\sigma - 1)/\sigma$ and $\alpha = \mu/\sigma$ (see Fujita et al., 1999, p. 54) the instantaneous equilibrium wage and price index in region $r$ can be written in a much simpler form. That is, the manufacturing wage for region $r$ is given by

$$w_r = \left[ \frac{Y_s \tau_{rs}^{1-\sigma}}{P_s^{1-\sigma}} + \frac{Y_r^{1-\sigma}}{P_r^{1-\sigma}} \right]^{\frac{1}{\sigma}} \tag{15}$$

and the price index becomes

$$P_r = \left[ \frac{1}{\mu} L_s^M (w_s \tau_{sr})^{1-\sigma} + \frac{1}{\mu} L_r w_r^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{16}$$

where $Y_s$ (or $Y_r$) is the income of consumers at region $s$ (or $r$). According to the previous assumptions, the wage of numéraire is equal to 1. We also assume that the share of manufacturing workers is $\mu$, and that of numéraire workers is $1-\mu$ in the whole economy and that each region is evenly endowed with numéraire workers. Then, if the share of manufacturing workers in region $r$ is denoted by $\lambda_r$, the income of region $r$ can be written as

$$Y_r = \mu \lambda_r w_r^M + 1 - \mu \tag{17}$$

Following the established tradition in NEG, in the short run it is assumed that markets can adjust instantaneously. Here the spatial allocation of manufacturing workers is considered as given. The wage of manufacturing workers, the price index and the workers' income of each are determined in the instantaneous equilibrium at a point in time.

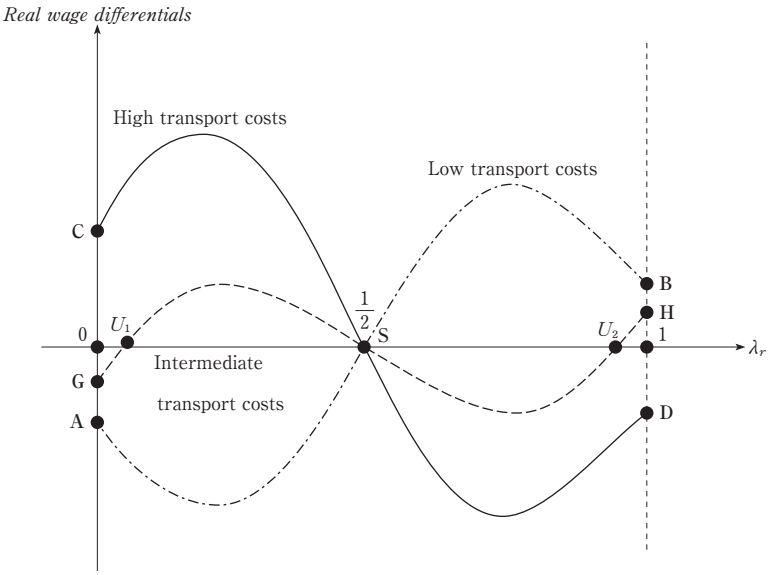
(4) **Long-run equilibrium**

In the previously described short-run equilibrium, the migration of skilled workers between regions is not considered, and the manufacturing distribution is fixed at one point in time. As time passes, the distribution of manufacturing workers will change across the two regions. Workers will move between them according to the indirect utility differential.

As the result of utility maximization, we can express the indirect utility of an industrial worker in region $r$ using income, manufacturing price index and price of numéraire goods as follows:

$$V_r = w_r P^{-\mu} p_A^{-(1-\mu)} \tag{18}$$

**Figure 1**: The wiggle diagram and local stability (from Baldwin et al., 2003, p. 26)




In the long run, skilled workers migrate between regions seeking higher indirect utility. As $p_A=1$ and the indirect utility ($V_r$) means the real wage $\omega_r$, which is the nominal wage deflated by the cost of living index $P^\mu$, we have

$$V_r=\omega_r=\frac{w_r}{P^\mu} \tag{19}$$

Just like Baldwin et al. (2003, p15), we can express the migration equation of skilled workers as

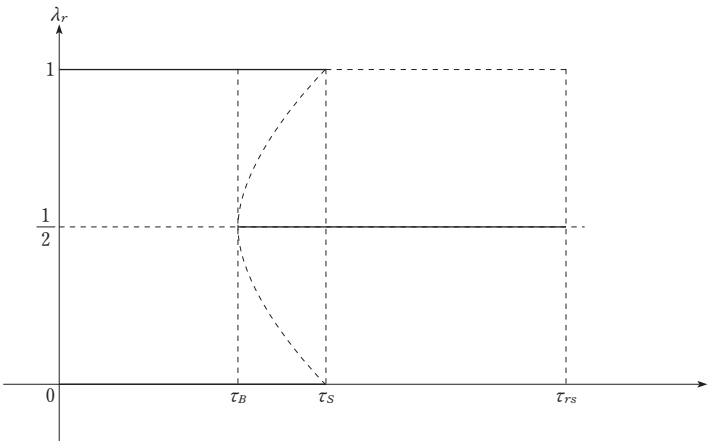
$$\dot{\lambda}_r=(\omega_r-\omega_s)\lambda_r(1-\lambda_r) \tag{20}$$

where $\lambda_r$ represents the share of skilled workers in region $r$.

The real wages are determined by a system of equations (15)–(17) and (19). Skilled workers can migrate between the two regions according to the difference of the real wages between them. Unfortunately, the simultaneous equations of (15)–(17) and (19) are too complicated for one to get the analytical solution, so numerical simulation is usually employed. The existing NEG models show how the decreases in transport cost affect the equilibrium distribution of manufacturing workers between the two regions, which can be demonstrated using the following two figures.

First of all, Figure 1 plots the relocation between real wage differentials of two regions and the share of manufacturing workers in region $r$ in accordance with different transport costs across regions. In the case of low transport costs , the curve of real wage differentials shows a unstable equilibrium point at S and the two stable ones at A and B, respectively. This means that all manufacturing workers would agglomerate to region $r$ or to another region $s$. In the case of high transport costs, the curve of real wages

**Figure 2**: The tomahawk diagram for Core-periphery model




differentials has only one stable equilibrium point S. This implies that both region $r$ and $s$ would have half of the whole manufacturing workers, i.e., industrial agglomeration would not occur. In the case of intermediate transport costs, there turns out to be three stable equilibrium points at $G$, $S$ and $H$, and two unstable ones at $U_1$ and $U_2$, respectively. This indicates that all manufacturing workers would agglomerate to region $r$ or region $s$, or the two regions would have half of manufacturing workers at the same time.

Secondly, Figure 2 shows the relation between the equilibrium share ($\lambda_r$) of manufacturing workers in region $r$ and transport costs ($\tau_{rs}$) across the region. From it, we can see that when transport costs are very high, the value of $\lambda_r$ would be 1/2, which implies that region $r$ and $s$ would share half of manufacturing workers simultaneously. When transport costs are very low, the value of $\lambda_r$ becomes 0 or 1. This means that all manufacturing workers would agglomerate to region $r$ or to region $s$. It can also be seen that when transport costs are intermediate, say between $\tau_B$ and $\tau_S$, the value of $\lambda_r$ could be 0, 1/2 or 1. That is, there would be a full agglomeration of manufacturing workers in region $r$ or region $s$, or there would be no agglomeration of them in either of the two regions. Here, $\tau_B$ and $\tau_S$ are called break point and sustain point, respectively, in NEG models.

Although the CP model is mathematically intractable, it conveys rich results and economic implications. Before the invention of analytical models, it enjoys popularity for many years. In the following, we review the refinements of the CP model.

### 2.2 The refinements of the CP model

Since the emergence of the CP model, the research issues of NEG have attracted concerns from mainstream economists. However, the problem of its mathematical intractability hampers its further progress. To date, the original CP model has relied heavily on the tricks that are called, in Fujita et al. (1999), "Dixit-Stiglitz, icebergs, evolution, and the computer." To avoid the handicap of CES setup with iceberg-form transport cost, people incorporate an alternative of utility and technological hypothesis and investigate the robustness of the results. Based on the thread of the original CP model, countless

refinements and extensions have been developed.

One of the most celebrated advances is the OTT model (Ottaviano et al., 2002). It steps out the original canonical framework, which incorporates distinctive consumer preference and transport costs. The merit of this model is mathematically tractable, which successfully incorporates the pro-competitive effect, i.e., decreasing profit-maximization in competition, which is a centrifugal force. Another distinction is that each consumer's spending on manufacturing varieties becomes independent of income, and the overlap originated from income effect disappears. But the OTT model still captures the main features of the original CP model.

Besides the OTT model, there are many other extensions on the basis of the CP model. They do not break away from the framework of "Cobb-Douglas-Dixit-Stiglitz-Iceberg" originally utilized in Krugman (1991). One is the footloose capital model, i.e. the FC model (Martin and Rogers, 1995). The FC model abandons many remarkable features of the CP model in order to obtain its tractability. For example, the mobile factor repatriates all of its earnings to its original region, and the causality of demand and the cost links are disappearing. But, the CP model's agglomeration mechanism is kept in the FC model, which is related to the home market effect that a region's economic activities encourage more industries to locate in the region. The merits of the FC model entail that it can be used to deal with many asymmetrical cases. However, it is also worth noting that, as an important feature of the FC model is the migration capital, it is more suitable for the study of capital movement between two countries.

The merits of the CP model are plain to see, but its intractability prevents it from further spreading and providing more economic policy guidance. So, seeking its analytical model becomes the main task of spatial economists. Forslid and Ottaviano (2003) modify a CP model and develop the solvable version of the CP model called the FE model. It assumes that both skilled and unskilled workers are employed in the manufacturing sector. Krugman and Venables (1995) and Fujita et al. (1999) extend the original CP model to express the migration of factors by the input-output linkage among firms. The parallel of forgoing FE and FC models are FCVL (i.e. Footloose Capital Vertical Linkage) (Robert-Nicoud, 2002) and FEVL (i.e. Footloose Entrepreneur Vertical Linkage) (Ottaviano, 2002). Those vertical linkage models show the same insights as the CP model, while FCVL and FEVL yield tractable solutions.

There is another parallel of these models. Most NEG models indicate that spatial agglomeration is the result of globalization and economic integration. But, when we detect the spatial evolution of some developing countries, as mentioned before, redispersion from the core region to the periphery can be observed. Especially in recent years, the patterns of spatial economics are much richer than those of the CP model. Many firms are relocated from the core to the periphery. In other words, industrial dispersion takes place in congested core regions or cities. By now, some papers have investigated this process under the NEG framework. Tabuchi (1998) and Tabuchi and Thisse (2002) synthesize the NEG framework with urban economics and show that urban congestion costs can bring

about redispersion of manufacturing with the decrease of transport costs. Picard and Zeng (2005) find that the agricultural sector can contribute to the redispersion of economic activities. Because of the requirement of agriculture labor in the manufacturing sector, two dispersion forces generated by the agriculture sector dominate the agglomeration forces. With further economic integration, relocation of manufacturing from the core to the hinterland is demonstrated. The changes of dispersion-agglomeration-redispersion seem to be a universal phenomenon. We need to employ new tools to explicitly explain the redispersion of spatial economy.

To date, the progress of NEG has been very fast. By using an alternative of utility functions and technological assumptions, NEG models have been improving. Nevertheless, this does not mean NEG has matured completely. When we review the existing models, we find that the FC model is more appropriate to investigate international trade, as it assumes that only goods but not workers are mobile between countries. Similarly, no production factors are mobile between countries in the FE model. But the international movement of capital such as FDI is an important feature of globalization. In this sense, NEG models have some application limits. Despite these drawbacks, we could still say the refinements of NEG models made so far have broadened and improved the explanation of spatial economics.

Since the emergence of CP model, many refinements have been made to explain the mechanism of economic agglomeration. Although these models show distinguished forms, their basic logic ideology still originates from Krugman's seminal paper. The models keep its typical features in their works and show many identical conclusions, but when we apply the alternative models of NEG, we find that the setting of two regions is unrealistic to explain the complex real spatial economy. In addition, as Ottaviano and Thisse (2004) point out, NEG allows one to better understand why agglomeration occurs; its models have little to say about where agglomeration arises. So it is needed to investigate why some regions are more successful than others. The setting with two regions is insufficient to explain the multi-regional system of the real economy. In most NEG models, the spatial economy is determined purely by economic mechanism.

The assumption of homogeneous space makes it easy to understand the mechanism of spatial agglomeration. However, there are some cases in which the geography (first nature) plays a more important role in shaping the landscape of spatial economy than the economic force (second nature). In fact, regions are never wholly symmetric with the same endowment, especially in terms of their trade accessibility to overseas markets and their scales.

## 3. From Symmetry to Asymmetry

As the foregoing discussion indicated, there are many situations in the real world in which the asymmetrical geographical features play a significant role in economic develop-

ment. In particular, in the spatial economy of the emerging economies, core regions are more successful than others. Their natural advantages are the indispensable element for them to become economic centers. It has been shown that China's industrial production presents strong agglomeration in its east coastal region, and the regional disparity between the interior and coastal region has been increasing (Fujita and Hu, 2001). Amiti and Javorcik (2008) show that access to customers and suppliers are the main determinants of FDI location in China.

There are many documents accounting for the impact of geographic handicaps on regional economic disparity. Gallup et al. (1999) find that geographic elements have major effects on economic growth through transportation costs and agriculture productivity. As Behrens and Gaigne (2006) discuss, some geographic elements, such as topography, climate and natural resources, are beyond the reach of economic policy. When we incorporate geographic features into our NEG models, the accessibility to main markets and regional sizes are important for economic consideration. Those assumptions of two symmetrical regions in early NEG models need to be modified. In this section, we provide a detailed summary about the related literatures. We divide these works into symmetrical and asymmetrical cases based on how they deal with the accessibility to markets and region sizes.

### 3.1 The symmetrical models

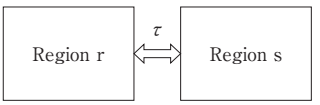
So far, there has been a growing number of theoretical literatures about the spatial distribution of economic activities among countries or regions in NEG. However, owing to their different springboards and hypotheses, there has not yet been a unified explanation of the distribution of economic activities. Many earlier papers focused only on the case of two regions or two countries as with trade theory and investigated the domestic or international disparities. Meanwhile, a few other works expand to include two countries or four regions. But, the regions and countries considered are still symmetrical in terms of access to markets and the size of regions and countries.

With respect to the earlier literature, symmetrical size and access of regions or countries are considered. Krugman (1991) initiates the classical CP model and endogenously reaps the economic agglomeration by utilizing monopolistic competition, increasing return to scale and iceberg-form transport costs. As a seminal paper, it inaugurates a new era, but the case of two symmetrical regions is investigated (see Figure 3).

Some other works focus on the industry agglomeration and relocation between countries (Martin and Rogers, 1995; Puga 1999, Puga and Venables 1996). They explore only the economic agglomeration or dispersion at the level of regions or countries independently and fail to analyzethem across regions and countries at the same time.

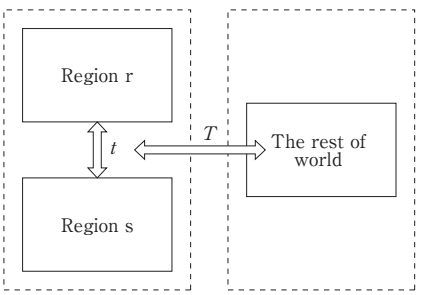
We note that some studies step forward and incorporate a setting of two domestic regions and consider the rest of world as the third region (see Figure 4). Krugman and Elizondo (1996) use such a setting to show that closed markets encourage regional convergence while open markets disaggregate it by considering the urban land rent and

**Figure 3**: The CP model



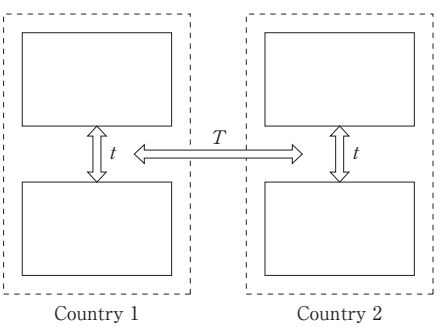

Notes: $\tau$ is the transport cost.

**Figure 4**: A model of two countries and three regions




Notes: $t$ is the transport cost, and $T$ is the trade cost.

**Figure 5**: A model of two countries and four regions



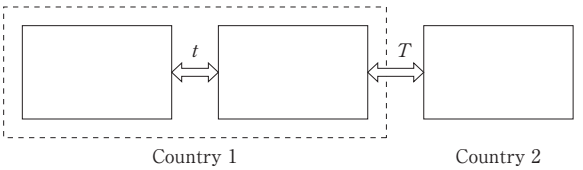

Notes: $t$ is the transport cost, and $T$ is the trade cost.

commuting cost as a centrifugal force. In contrast, assuming that immobile workers are a centrifugal force, Paluzie (2001) reaches the opposite conclusion that trade liberalization exacerbates the regional inequalities within a country. In these models, although the impact of international trade on industrial location is considered, the two domestic regions are assumed to be endowed with identical size and access to the world market.

In a paper investigating spatial economy, Behrens (2011) incorporates two countries involving three regions, and finds a complex relationship between regional disparity and international trade costs and domestic transport cost. In his paper, two countries are considered, but regions are still symmetrical in terms of their sizes and accessibility to the world market.

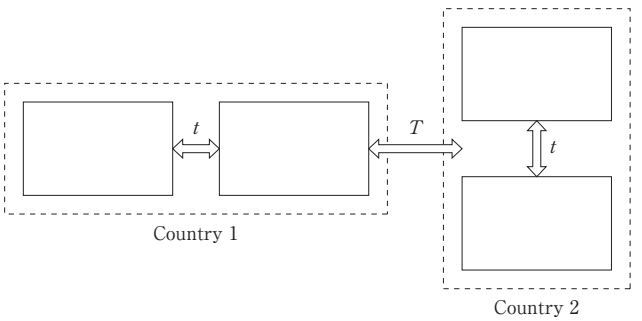
A few other researchers present a framework having two countries and four regions (see Figure 5). Monfort and Nicolini (2000) analyze such a framework and conclude that the economic integration of countries results in the emergence of regional economic

**Figure 6**: The asymmetrical model of three regions




Notes: $t$ is the transport cost, and $T$ is the trade cost.

**Figure 7**: The asymmetrical model of two countries and four regions (1)




Notes: $t$ is the transport cost, and $T$ is the trade cost.

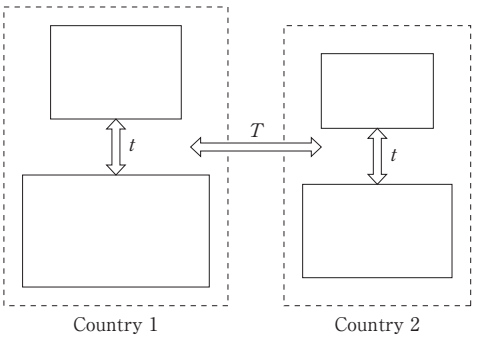
agglomeration. Behrens et al. (2006b, 2007) successfully obtain an analytical result for such a model of two countries involving four regions using the quadratic utility function and linear transport costs. But in their papers, the domestic transport costs of the two countries are assumed to be equal. It is unrealistic that the two countries, such as a developed country and a developing one, have the same infrastructure level.

### 3.2 The asymmetrical models

Unlike the above-mentioned symmetrical models, some recent papers also pay much attention to geographical asymmetry by assuming a border region or hub. Ago et al. (2006) analyze the evolution of spatial economy in a model of three regions located on a line with a decline of transport cost. They show that the central region always has a location advantage, but it will weaken or even disappear by using a quadratic-linear function. However, there is no international trade considered in their model, and the regional sizes are equally assumed. Crozet et al. (2004) assume a border region in their model involving two domestic regions and the rest of the world, and demonstrate that domestic regional agglomeration depends only on the trade liberalization (Figure 6). But in their paper, the effects of domestic transport costs are neglected. Behrens et al. (2006a) investigate the impacts of changes in domestic transport costs and international trade on regional industrial location (Figure 7). The regions in a country have asymmetrical accessibility to overseas markets, but their size remains symmetrical.

Finally, Zeng and Zhao (2010) analyze the relation between the interregional (international) inequalities and international trade cost and domestic transport costs by assuming

**Figure 8**: The asymmetrical model of two countries and four regions (2)




Notes: $t$ is the transport cost, and $T$ is the trade cost.

asymmetrical sizes of countries and regions (Figure 8). Unfortunately, in their paper, regional accessibilities to overseas markets are still symmetrical.

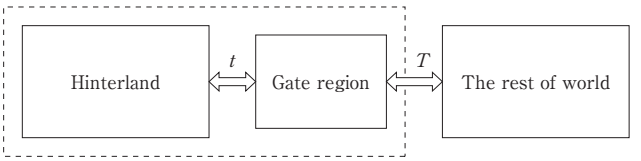
## 4. Toward Full Asymmetry

As explained, NEG has reached a mature phase, especially from the viewpoint of the developments of its two-region models. However, these models are still unable to explain the complex and asymmetrical spatial economy. NEG offers a satisfactory explanation for why spatial agglomeration happens. But it has explained little about where such an agglomeration emerges. Its existing symmetrical models indicate that spatial agglomeration always emerges in one of the two regions. As Behrens et al. (2006b) point out, most NEG models expunge the geography by confirming themselves to "double-point economics." In fact, a tiny change of accessibility of domestic regions produces more abundant results than the basic CP model. Besides the asymmetry of accessibility, when size of regions and countries are assumed to be asymmetric, the region disparity can exhibit an inverted U-shape with respect to trade costs in the FC framework (Zeng and Zhao, 2010). Existing research has not studied cases in which both regional sizes and accessibilities to overseas market are asymmetrical.

In the real world, the hierarchy of space economy is characterized by different accessibilities and sizes. In China, for example, coastal regions have better access to international markets, while more unskilled agriculture workers live in the hinterlands. This applies to other emerging economies in the world, as well. It is important to incorporate such a full asymmetry into the traditional NEG models to explain complex spatial economy.

As an attempt to accurately reveal the reality of developing countries like China, we could consider the home country has two asymmetrical regions. The hinterland region has to access to overseas markets through the gate region, and the former is also endowed with more unskilled workers than the latter (Figure 9). We believe that such an incorporation of regional asymmetry could contribute to the exiting theoretical models of

**Figure 9**: A model of full asymmetrical regions




Notes: $t$ is the transport cost, and $T$ is the trade cost.

NEG.

Forslid and Ottaviano (2003) present a model having asymmetrical regional sizes and find more abundant results than the symmetric CP model. As a result, the region with more immobile workers attracts a large number of manufacturing firms. This is because the centrifugal forces in NEG models are mainly the demands of immobile workers. In the reduction of transport costs, the region having many of those workers shows the advantage of attracting more manufacturing firms. Such an asymmetrical situation seems to be more prevalent in the developing world, especially in BRICs countries.

Furthermore, Behrens et al. (2006b) show that when the domestic regions are well integrated, firms prefer to locate in a region that has better access to overseas markets. It would be interesting to see what would happen if the hinterland were endowed with more immobile workers while the gated region had better accessibility. When we allow the hinterland to have different sizes, more interesting results may arise. Besides these asymmetrical considerations, to better explain the reality of developing countries, the agriculture sector is also worth being incorporated into NEG. It can be seen that unskilled workers are employed not only in the agriculture sector but also in the manufacturing one. These considerations could enrich the centrifugal forces in the interaction of economic activities. We hope to offer a better explanation about why spatial agglomeration happens and where it arises.

In the past three decades, China has experienced a rapid urbanization and an increasing regional disparity in globalization. But in recent years, dispersion of manufacturing industries is often mentioned by the media. In addition, the economic growth rates in the central and western provinces have become higher than eastern coast regions. Such a process of industrial dispersion attracts more attention from local governments and scholars. Regarding such a transition from agglomeration to dispersion, we need to give more explicit and persuasive explanation. At present, few works have been done to investigate such a dispersion process. It is obvious that the objective of firm migration is the seeking of a better location where the prices of labor are cheaper. This process also confirms the importance of comparative advantages, which means that to consider the full asymmetry would also help us to explore the recent industrial dispersion in China.

## 5. Concluding Remarks

NEG has been developing for two decades since the birth of Krugman's Core-Periphery model. It has conceptually matured and attracted much attention from mainstream economics. NEG employs the pure economic theories and successfully explains the endogenous formation of spatial agglomeration. But one of its weak points is geography. It remains unable to explain the real and complex hierarchy that characterizes the spatial economy. When we observe the space of the emerging economies, some core regions are more successful than others. The natural advantages of these regions are indispensable to their being economic centers. It is important to consider these natural advantages such as regional accessibilities and sizes in the explanation of spatial economy.

This paper presented an overview on the recent developments of NEG and proposed some ideas about further research. Based on a retrospect of the developments of NEG, we found that many existing models are isomorphic, showing identical results in a similar way. The existing two-region model is very useful, but there is still a gap between theory and reality. For this reason, we think that it is needed to consider the full asymmetry in terms of regional accessibilities and sizes in NEG. We hope to use this approach to better explain the recent dispersion of manufacturing activities in many emerging countries in the near future.

### Notes

1) One variety only is produced in one location, so the number of firms equals the number of variety.

### References

Ago, T., Isono, I., Tabuchi, T., 2006. Locational disadvantage of the hub, *Annals of Regional Science* 40, 819–848.

Amiti, M., Javorcik, B. S., 2008. Trade costs and the location of foreign firms in China, *Journal of Development Economics* 85, 129–149.

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., Robert-Nicoud, F., 2003. *Economic Geography and Public Policy*. Princeton University Press.

Behrens, K., 2011. International integration and regional inequalities: how important is national infrastructure? *The Manchester School* 79(5), 952–971.

Behrens, K., Gaigne, C., 2006. Developing the outermost regions of Europe: Some lessons from economic geography. Report prepared for the conference: La situation macroeconomique de l'ultraperipherir europeenne. Brussels, December 15, 2006.

Behrens, K., Gaigne, C., Ottaviano, G., Thisse, J-F., 2006a. Is remoteness a locational disadvantage? *Journal of Economic Geography* 6, 347–368.

Behrens, K., Gaigne, C., Ottaviano, G., Thisse, J.-F., 2006b. How density economies in international transportation link the internal geography of trading partners. *Journal of Urban Economics* 60, 248–263.

Behrens, K., Gaigne, C., Ottaviano, G., Thisse, J.-F., 2007. Countries, regions and trade: on the welfare impacts of economic integration. *European Economic Review* 51, 1277–1301.

Crozet, M., Koenig Soubeyran, P., 2004. EU Enlargement and the Internal Geography of Countries. *Journal of Comparative Economics* 32, 265–279.

Dixit, A. K., Stiglitz, J. E., 1977. Monopolistic competition and optimal product diversity. *American Economic Review* 67, 297–308.

Fally, T., R. Paillacar, and C. Terra., 2010. Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics* 91, 155–168.

Fan, C. C., Sun, M., 2008. Regional Inequality in China, 1978–2006. *Eurasian Geography and Economics* 49(1), 1–20.

Forslid, R., Ottaviano, G. I. P., 2003. An analytically solvable core-periphery model. *Journal of Economic Geography* 3, 229–240.

Fujita, M., Krugman, P. R., Venables, A. J., 1999. *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge, Massachusetts.

Fujita, M., Hu, D., 2001. Regional disparity in China 1985–1994: the effects of globalization and economic liberalization. *The Annals of Regional Science* 35, 3–37.

Fujita, M., Mori, T., 2005. Frontiers of the new economic geography. *Papers in Regional Science* 84 (3), 377–407.

Gallup, J. L., Sachs, J. D., Mellinger, A., 1999. Geography and economic development, *International Regional Science Review* 22, 179–232.

Golubchikov, Oleg., 2006. Interurban Development and Economic Disparities in a Russian Province, *Eurasian Geography and Economics* 47(4), 478–495.

Hanson, G. H., 1996. Agglomeration, dispersion, and the pioneer firm. *Journal of Urban Economics* 39, 255–281.

Henderson, J. V., Lee, T., Lee, J.-Y., 2001. Scale externalities in Korea. *Journal of Urban Economics* 49, 479–504.

Krugman, P. R., 1991. Increasing returns and economic geography. *Journal of Political Economy* 99, 483–499.

Krugman, P. R., 2011. The new economic geography, now middle aged. *Regional Studies* 45(1), 1–7.

Krugman, P. R., Venables, A. J., 1995. Globalization and inequality of nations. *Quarterly Journal of Economics* 110(4), 857–880.

Krugman, P., Elizondo, R. L., 1996. Trade policy and the third world metropolis, *Journal of Development Economics* 49, 137–150.

Martin, P., Rogers, C. A., 1995. Industrial location and public infrastructure. *Journal of International Economics* 39, 335–351.

Monfort, P., Nicolini, R., 2000. Regional convergence and international integration. *Journal of Urban Economics* 48, 286–306.

Ottaviano, G. I. P., 2002. Models of 'new economic geography': factor mobility vs. vertical linkages. Mimeo, Graduate Institute of International Studies.

Ottaviano, G. I. P., Tabuchi, T., Thisse, J.-F., 2002. Agglomeration and trade revisited. *International Economic Review* 43, 409–436.

Ottaviano, G., Thisse, J.-F., 2004. Agglomeration and economic geography. In: Henderson, J.V., Thisse, J.-F. (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4. Cities and Geography. North-Holland, Amsterdam, pp. 2563–2608.

Paluzie, E., 2001. Trade policy and regional inequalities. *Papers in Regional Science* 80, 67–85.

Picard, P., Zeng, D.-Z., 2005. Agricultural sector and industrial agglomeration. *Journal of Development Economics* 77, 75–106.

Puga, D., 1999. The rise and fall of regional inequalities. *European Economic Review* 43, 303–334.

Puga, D., Venables, A. J., 1996. The spread of industry: Spatial agglomeration in economic development. *Journal of the Japanese and International Economies* 10, 440–464.

Robert-Nicoud, F., 2002. A simple geography model with linkages and capital mobility. Mimeo, London School of Economics.

Robert-Nicoud, F., 2005. The structure of simple 'new economic geography' models (or, on identical twins). *Journal of Economic Geography* 5(2), 201–234.

Tabuchi, T., 1998. Agglomeration and dispersion: a synthesis of Alonso and Krugman. *Journal of Urban Economics* 44, 333–351.

Tabuchi, T., Thisse, J.-F., 2002. Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics* 69, 155–177.

Zeng D.-Z., Zhao L.-X., 2010. Globalization, interregional and international inequalities. *Journal of Urban Economics* 67, 352–361.