□論 文□

Agglomeration Economies, Trade, And System of Cities : A General Equilibrium Approach

권 영 각
權 寧 珏
(交通開發硏究院 先任硏究員)

目 次

요 약

본 연구는 도시규모 결정에 관해 수요 또는 공급측면만을 중시하는 재래식 접근 방법을 통합, 개선하는 일반균형의 모형을 제시하고 이에 따라 국가 도시규모체계의 효율성을 이해하는데 그 목적이 있다. 이 모형의 주요한 요소는 도시집적이익 및 불이익, 산업구조 그리고 도시간 자원이동 및 무역을 통한 상호의존성등이다.

엄밀한 이론적, 절대적 적정 도시규모체계는 모든 도시가 완전자립하에 단위도시 적정규모를 이루어 동일규모일때 가능하며 시민의 복지가 극대화된다. 그러나 실제 인적, 물적 이동성이 완벽하지 않은 현실하에서는 도시규모간 위계성이 생기게 되며 이는 도시간 무역을 가능케하여 상대적 도시규모의 적정성을 대변해 준다.

Ⅰ. INTRODUCTION

According to recent UN reports, eighty-three percent of 114 developing countries indicated that a heavy concentration of population in a few large metropolitan centers was unacceptable (UN[1978, 1980]). The emerging concensus among them is that the uneven spatial distribution of population is the main obstacle to socioeconomic development. To the countries, the existence of their largest cities is reminiscent of economic inefficiency, regional inequality, political unrest, and the familiar "big city headaches" such as congestion, pollution, and overburdened infrastructure and housing.

In practice, therefore, most urban policies have mainly been limited to alleviating the situation in their biggest cities, either to direct more public investments into them or to divert people away from them. Rarely has there been a serious commitment to develop the biggest city along with smaller cities within the context of the national system of cities. what is lacking is the view that the unit of national urban policies must be the national economy rather than the city economy. However inadequate the current urban policies may be, they nevertheless seem to reflect the current state of knowledge in the literature. Although stylized facts such as the rank-size rule and paradigms such as "concentrated decentralization"(Rodwin[1961], Alonso[1968]) exist, it is surprizing that no study has analytically dealt with the basic issue : what spatial allocation of populations among cities is commensurate with national economic development.

This dissertation attempts to address the issue by devising an analytic spatial economic framework in which the economic efficiency and welfare of a city depend not only on its own size and internal characteristics, but also on those of other cities within the nation. It is hoped that this framework can serve as an analytic basis of national urban policies in enhancing the overall level of welfare. To this end, we will develop a spatial general equilibrium urban model in which cities are viewed as centers of production, consumption, and trade. Essential elements of the model will include sizes of cities, industrial composition and production technologies, and economic interactions of cities via migration and trade.

II. MODELS OF CITY SIZE AND OUTPUT DETERMINATION

The standard practice regarding the determination of city size and output is to equate city size with city laborers. Since the number of laborers is highly dependent on net migration into a city, an emphasis is put on the relationship between employment and migration. Two contrasting approaches, which emphasize either the demand or supply side of the labor market, will be reviewed.

1. Demand-and Supply-oriented Models

Muth[1971], Engle[1974], Schaefer [1977], and Miron[1979] provide a useful classification of urban models on city size and output : the demand-and supply-oriented approach. The former type, or the external approach, inherits the Keynesian heritage in that the main emphasis is on exogenous demands from other areas. Factor supplies are presumed to be perfectly elastic so that factor in-migrations to the city are limited only by the city's factor demands at given real factor prices. It is further assumed that the demand for the product is price inelastic and that no matter how much is demanded externally, it will be supplied by the city. Spatially, this implies that the size and output of the city are wholly determined by its output market areas. Migration becomes the consequence of the city's growth rather than the cause of it. The export-base model and the various urban hierarchy models, such as Losch[1954], Beckmann[1958], and Beckmann and McPherson[1970], belong to this type.

Although there is a consideration in this approach of interdependence among cities via exports and factor exchange, it usually does not take account of the city's internal production and consumption characteristics. With no explicit consideration of comparative costs in production, the level of exports is often exogenously determined, but there is a complete lack of explanation as to how the level of imports is determined. This approach essentially envisions a hierarchy of cities based on trade imbalance. Goods are exported only down the hierarchy from larger to smaller cities, but there are no corresponding reverse flows.

Representing the tradition in the trade literature, the supply-oriented or internal approach is concerned with one city in isolation, and focuses on the city's production technology, endowment of factors, and other internal characteristics. It is often assumed that factor demands are perfectly price elastic, but factor supplies are inelastic due to non-wage considerations. Further, since the city is presumed to face a perfectly elastic export demand at a given output price, it is suggested that no matter how much the city produces, it will be purchased elsewhere. Consequently, the output and size of the city are limited only by the internal cost consideration and available factor supplies. Migration now becomes the cause of the city's growth, rather than the response of it.

In this approach, an exogenous shock, for example, to shift the demand for labor will induce migration into the city, and this will in turn result in an even greater equilibrium city size and output(Borts and Stein [1964], Thompson[1968], and Muth[1971]). Although this approach emphasizes the city's various internal characteristics, its neglect of other cities as both output market and factor supply areas make it highly unrealistic as a basis of a spatial urban model. However, since this approach has been influential and pervasive, we will now look into it further.

2. Localization and Urbanization Economies

Following the Heckscher-Ohlin practice, proponents of the pure supply-oriented approach try to find the raison d'etre of the city in its resource endowment that gives rise to its comparative advantage. However, the approach is generally extended to contend that when a city is assumed to be built on homogeneous land, the main economic justification for the city is the presence of agglomeration economies in production and/or consumption. After Ohlin [1933], Hoover[1937], and Isard[1956], it is widely accepted that agglomeration economies can be realized at the firm (scale economies), the industry (localization economies), and the city level(urbanization economies). In the urban context, however the last two types of agglomeration economies, which are external to the firm, get the most attention. Under external economies, firms are usually modelled to operate under constant returns to scale and to behave competitively(Kemp[1955, 1964], Melvin [1969], Chipman[1970]).

Mills[1967] was among the first to present a supply-oriented model that incorporates urbanization economies in the production function of a composite urban good. In his monocentric city where firms are located in the city center surrounded by residential areas, a unique equilibrium size is reached by a sharp increase of commuting costs and congestion(urbanization diseconomies) after certain city sizes are attained. The equilibrium city size represents an optimum at which net urbanization economies(urbanization economies less diseconomies) are maximixed. The equilibrium is essentially an autarkic one, and the optimum city size is determined entirely by the internal conditions. However, the equilibrium size should mean either of two paradoxical results : there is one city in the country ; or all cities are identical. Of course, the naivety of one universal optimum city size is well known.

One solution to avoid the paradox is Henderson's recognition that different-sized monocentric cities perform different functions, and they may operate at different, but efficient, equilibium sizes(Henderson [1974]). Each city, producing one distinct export good, is engaged in free trade with other-type (or size) cities at exogenously determined terms of trade. Thus, multiples of different-sized cities can coexist in equilibrium ; however, cities of the same type must be identical in size. It should be noted that Henderson's model is based on complete specialization of cities in which no distinction can be drawn between the city and the industry. Accordingly, a city, however big it may be, must always be engaged in trade, and unlike the standard trade model, factor movements are not a substitute for trade.

The complete specialization is the result of Henderson's a priori reliance on the so-called small city approximation in which the city produces (and consumes) on such a small scale that the effect of its production (and consumption) on the national (and international) markets can be ignored. Although complete specialization is always a possibility, the actual degree of specialization is to be limited by the extent of the market. In light of the fact that the contemporary urban concern stems mainly from the alleged gigantism of the largest cities and their pervasive effects on national economies, the small city approximation seems particularly inappropriate.

The possibility of multiple industries in the city is more a rule than an exception, however. This is because, to the extent that the marginal products of factors are finite and diminishing, greater specialization by the city in a particular good will eventually increase the marginal (social) opportunity cost of that good. Moreover, as the "home market effect" of the trade literature suggests, it can be argued that the city's size of the internal market determines the range of goods produced, or industrial diversity. Indeed, industrial diversity or "breadth", to quote Thompson[1968], must be a fundamental characteristic of the city.

3. Synthesis

A truly general spatial urban model calls for an integration of both approaches. In particular, there is a need to incorporate into our model some refinement of the production conditions typified by the supply—oriented approach, as well as the consideration of demand for the city's output of both internal and external origin indicated by the demand—oriented approach. Once we introduce two industries into our model in order to allow for industrial diversity, there remains the problem of specifying urbanization and localization economies. Despite deliberations made by Hoover and Isard, however, the relatinoship between the two agglomeration economies does not seem well established. Consequently, we will handle this problem by the following set of assumptions.

First, we envision the existence of a single homogeneous city labor market from which labor is allocated between the two industries. The size increase of the labor market along with its greater spatial concentration would enhance its efficiency and allow for a greater realization of localization economies for both industries. Second, instead of introducing urbanization economies separately, we assume that they are the result of localization economies. Thus, urbanization economies are presumed to consist of localization economies only. Although this assumption excludes the possibility of inter—industry production exter nalities, this nevertheless generates a conceptually more sound measure of urbanization economies than most measures based on the city size alone. The measure will be in utility terms, and it will reflect not only the city size but also the underlying industrial composition of the city in equilibrium.

Ⅲ. A SINGLE—CITY MODEL

1. Introduction

The purpose of this chapter is to formulate a spatial general equilibrium model of a single city under autarky. With urbanization economies and diseconomies, we will show how a unique equilibrium is obtained under perfect competition. In the end, factors, output, their prices, and the level of welfare are all determined endogenously, given the resource endowment and the internal production and consumption characteristics. The fact that the equilibrium is attained totally by the internal characteristics puts this single—city model in the tradition of the supply—oriented approach. Later, we will allow two cities of the nation to be engaged in factor exchange but not in trade. This will present us with a city size distribution which reflects the extension of the supply—oriented approach.

The representative city is based on following characteristies.

(1) The city is meant to be : a homogeneous monocentric area the center of which is occupied by the point Central Business District (CBD) for firms ; surrounded by residential area ; and economically the place where its people live and work.

(2) It "produces" housing or land, for residential use only, and two goods for own

consumption and possible interurban trade.

(3) It is spatialin that housing is differentiated by its distance from the work center (CBD), and its rents are determined accordingly. The the two goods are not subject to transport costs both within and between cities.

(4) Residents are identical in all respects such as skills, utility functions, vocational preferences, and capital ownership. The city land is collectively owned, and each has an equal share (a la land bank). Thus, total housing rents are to be equally divided among residents.

(5) The two industries employ two homogeneous factors, labor and capital, which are fully employed and mobile.

(6) All firms in one industry are identical in size, and share an identical production function, but those between industries are different.

Being a localized nontrdable good, the prices of housing and the associated commuting costs are uniquely determined within the city. Because the increase of city size would push up the costs, they will be the measure of urbanization diseconomies in our model.

2. Consumer Equilibrium

There are N residents in the city with the individual utility function, u, which incorporates the fixed housing consumption set to unity and the fixed demand for leisure :

$$u=x_1^{\theta_1}x_2^{\theta_2}, \ \theta_1, \ \theta_2>0, \ \theta_1+\theta_2=1, \quad (3.1)$$

where xi=individual consumption of good i, (i=1,2). We normalize to unity the fixed amount of total available time after leisure. The representative resident then will allocate it between work and commuting as in the following :

$$\ell(t)+gt^2=1, \quad 0\leq t\leq m, \quad (3.2)$$

where $\ell(t)$=labor supply at location t,

g = a technologically determined commuting parameter,

gt^2=time spent for commuting at location t,

m =distance of the city boundary from the CBD.

Consequently, although the demand for leisure is fixed, the supply of labor is variable with respect to location. In light of congestion, the commuting time function, gt^2, is based on a premise that the marginal commuting time is increasing with respect to distance. At location t, commuting costs the resident the foregone wage of wgt^2 where w is the wage rate : only the time aspect of commuting is considered. Thus, the economic consequence of commmuting, or of urbanization diseconomies, is a waste of some of the total available labor for the city as a whole.

The utility function is to be maximized subject to the budget constraint :

$$p_1 x_2+p_2 x_2+p_h(t)-w\,\ell(t)-I=0, \quad (3.3)$$

where pi=price of good i (i=1,2),

$p_h(t)$=price of housing service (rent) at location t,

I =nonwage income.

Aside from location, residents will set the marginal utilities of the two goods proportionally to respective prices, yielding the following individual demand functions :

$$p_i x_i = \theta_i [w \ell(t) + I - p_h(t)]. \quad (3.4)$$

Substituting (3.2) and (3.4) into (3.1), we get the indirect utility function,

$$V = (\theta_1/p_1)^{\theta_1} (\theta_2/p_2)^{\theta_2} [w(1-gt^2) + I - p_h(t)]. \quad (3.5)$$

Spatial equilibrium in the housing market requires that $\partial v / a \partial t = 0$. Differentiating (3.5), it turns out that

$$\partial p_h(t) / \partial t = -2wgt. \quad (3.6)$$

At any location, foregone wages due to increased commuting must just be offset by reduced rents, thereby leaving everyone indifferent as to location. Integrating the above, with the added assumption that $p_h(m) = 0$, or at the city edge the urban rent is zero, we can derive the rent gradient function,

$$p_h(t) = wg(m^2 - t^2). \quad (3.7)$$

With per capita housing consumption fixed at unity, the city boundary will be determined by $\pi m^2 = N$, where N=population of the city. We can rewrite (3−7) that clears the housing market :

$$p_h(t) = wg(N/\pi - t^2). \quad (3.8)$$

We are ready to calculate per capita nowage income, I. First, there is the rent income origination from the equal housing share provision. In our monocentric city, total housing rents are defined as

$$R = 2\pi \int_0^m t p_h(t) dt. \quad (3.9)$$

From the above, R and the per capita amount, R/N, become

$$R = \frac{wgN^2}{2\pi}, \qquad \frac{R}{N} = \frac{wgN}{2\pi}. \quad (3.10)$$

The remaining nonwage income comes from capital rentals, and per capita capital rental income is r(K/N), where r=capital rental, K=capital stock in the city. It is to be emphasized that with variable labor supply, K/N denotes per capita capital, not the usual capital−labor ratio of the city. Thus, per person nonwage income is

$$I = \frac{wgN}{2\pi} + \frac{rK}{N}. \quad (3.11)$$

Substituting (3.2), and (3.11) into (3.4), and aggregating it across the N individuals, we can derive the market demand functions as

$$p_i x_i = w\theta_i N \left(1 - \frac{gN}{2\pi} + \frac{rK}{wN}\right) (i=1,2), \quad (3.12)$$

where x_i=total demand for good i. we now have obvious consumer equilibrium conditions : the proportion of the resident's income allocated to good i is equal to θ_i.

The same substitution into (3.5) leads to an updated indirect utility function :

$$v = \theta_1^{\theta_1} \theta_2^{\theta_2} \left(\frac{w}{p_1}\right)^{\theta_1} \left(\frac{w}{p_2}\right)^{\theta_2} \left(1 - \frac{gN}{2\pi} + \frac{rK}{wN}\right). \quad (3.13)$$

Regardless of location, therefore, individuals derive the same level of welfare. With K/N fixed, utility becomes a linearly decreasing function of N. This is what we have expected with the introduction of urbanization diseconomies alone.

We now proceed to calculate the city−wide labor supply, L.

$$L = 2\pi \int_0^m t \ell(t) dt, \quad (3.14)$$

which from (3.2) is to be

$$L=N-\frac{gN^2}{2\pi}. \qquad (3.14)$$

Out of the total availabl3 time after leisure, N, $gN^2/2\pi$ is spent on commuting, leaving the city with the above for L.

The labor supply function (3.15) is a quadratic function the relevant portion of which is limited to the rising part of it. The upper bound of N is entirely determined by the situation of the person at the city edge. That is, he must be able to trade off the lower housing rent at that location against the increased commuting time and the consequent reduction in wage. Two behavioral assumptions are in order. First, on a daily basis each person sets asides 12 hours for leisure and another for work and commuting the latter of which is normalized to unity. Second, the round-trip commuting time for the resident at the edge is likely to be limited to no more than 6 hours. From (3.2) and (3.15), this means

$$0\leqq N\leqq\frac{\pi}{2g}, \quad 0\leqq L\leqq\frac{3\pi}{8g}. \qquad (3.16)$$

Therefore, if cities are endowed with equ al K/N, large cities will have higher capital-labor ratios, K/L, than smaller ones. This is the direct consequence of urbanization diseconomies.

3. Producer Equilibrium

In the city, there are two industries, each producing a distinct good. Given that all firms within each industry are identical in size and technology, the firm production functions can be aggregated into the industry production function,

$$Y_i=\lambda_i L_i^{\alpha_i} K_i^{1-\alpha_i}, \quad 0<\alpha_i<1, \quad 0<\lambda_i, \ (i=1,2), \qquad (3.17)$$

where : Y_i=output of industry i,

L_i, K_i=labor, capital of industry i.

Localization economies, or what Chipman[1970] calls "parametric external economies of scale", are introduced by λ_i. In this specification, λ_i is viewed by the firm as a Hicks neutral shift factor in making its business decisions. Because the rest of (3.17) constitute a homogeneous-of-degree-one production function, the collective behavior of the firms assures that the industries have constant-returns-to-scale production functions. This guarantees the exhaustion of total revenues by factor payments. Thus the existence of localization economies is consistent with perfect competitition.

Although the term λ_i is external to the firm, it is internal to the industry, and is actually related to the i-th industry's output :

$$\lambda_i=Y_i^{\varepsilon_i}, \quad 0<\varepsilon_i<1, \quad (i=1,2). \qquad (3.18)$$

When (3.18) is substituted into (3.17), we obtain

$$Y_i^{1-\varepsilon_i}=L_i^{\alpha_i} K_i^{1-\alpha_i}, \text{ or } Y_i=\left(L_i^{\alpha_i} K_i^{1-\alpha_i}\right)^{\rho_i},$$

$$\rho_i=\frac{1}{1-\varepsilon_i}>1, \ (i=1,2). \qquad (3.19)$$

The industry production function of (3.19) is homogeneous of degree ρ_i. With $\rho_i>1$, it will actually exhibit increasing returns to scale. According to Chipman, this is called the "objective" production function, whereas (3.17) is called the "subjective" production function. The essential difference is that while the former is based on actual production properties, the latter is based on entrepreneural behavior. Not only is the term λ_i external to the firms, but the relation (3.18) is assumed to be unknown to them. The determination of competitive equilibrium under localization economies requires the use of both.

According to the first-order conditions for profit maximization, the reward to each factor is the value of its marginal product to the firm, not to the industry. For such an entrepreneural decision, we compute the marginal private, or "subjective" products from (3.17), holding λ_i constant:

$$\frac{\partial Y_i}{\partial L_i}=\lambda_i\alpha_i\left(\frac{K_i}{L_i}\right)^{1-\alpha_i},\ \frac{\partial Y_i}{\partial K_i}=\lambda_i(1-\alpha_i)\left(\frac{K_i}{L_i}\right)^{-\alpha_i} \tag{3.20}$$

Substituting (3.17) into (3.20), we obtain

$$\frac{\partial Y_i}{\partial L_i}=\alpha_i L_i^{\rho_i-1}\left(\frac{K_i}{L_i}\right)^{\rho_i(1-\alpha_i)}=\alpha_i\left(\frac{Y_i}{L_i}\right),$$

$$\frac{\partial Y_i}{\partial K_i}=(1-\alpha_i)L_i^{\rho_i-1}\left(\frac{K_i}{L_i}\right)^{\rho_i(1-\alpha_i)^{-1}}$$

$$=(1-\alpha_i)\left(\frac{Y_i}{K_i}\right). \tag{3.21}$$

It is to be noted that the above marginal private products become less than the marginal social or "objective" products below, which can be computed directly from (3.19), by a factor of ρ_i:

$$\frac{\partial Y_i}{\partial Y_i}=\rho_i\alpha_i\left(\frac{Y_i}{L_i}\right),\ \frac{\partial Y_i}{\partial K_i}=\rho_i(1-\alpha_i)\left(\frac{Y_i}{K_i}\right). \tag{3.22}$$

Naturally, we assume that all marginal products are positive, finite, diminishing, and smaller than the average products. From (3.19), (3.21), and (3.22), this requires additional constraints on ρ_i as follows:

$$\rho_i\alpha_i<1, \qquad \rho_i(1-\alpha_i)<1. \tag{3.23}$$

Under perfect competition, a uniform wage rate, w, and capital rental, r, must prevail between the industries. From (3.21), therefore, we arrive at the following producer equilibrium conditions in which the factor payments exactly meet the outputs:

$$L_i=\alpha_i\left(\frac{p_i}{w}\right)Y_i,\quad K_i=(1-\alpha_i)\left(\frac{p_i}{r}\right)Y_i. \tag{3.24}$$

The above can be used to reveal the actual production properties via the marginal private cost curve, or the industry supply curve. Combining the objective production function (3.19) and (3.24), we obtain

$$\frac{p_i}{w}=\alpha_i^{-\alpha_i}(1-\alpha_i)^{\alpha_i-1}\left(\frac{r}{w}\right)^{1-\alpha_i}Y_i^{-\varepsilon_i},$$

$$\varepsilon_i=\frac{\rho_i-1}{\rho_i}. \tag{3.25}$$

The supply curve has constant elasticity $\rho_i/(1-\rho_i)$. Because $\rho_i>1$, the supply curve is negative-sloping. The effect of localization economies is that at a given factor price ratio w/r, the supply price relative to the wage rate, p_i/w, is decreasing with respect to the output.

The above profit-maximizing decisions, however, also represent a socially efficient production schedule. This is because the private marginal rates of technical substitution calculated from (3.21) are identical to the social ones from (3.22):

$$\frac{\partial Y_i/\partial L_i}{\partial Y_i/\partial k_i}=\frac{\alpha_i}{1-\alpha_i}\ \frac{K_i}{L_i}. \tag{3.26}$$

Despite the presence of localization economies, production is efficient, and is operated along the city's production possibility curve.

Although production is efficient in the sense of the identical marginal rates of factor substitution for the two industries, the marginal rate of product substitution in production (or marginal rate of transformation) is generally not equal to the price ratio. In order to avoid this analytic complication, we limit ourselves to the kind of localization economies which ensure equality between the marginal rate of transformation and the price ratio, by adding the constraint, $\rho = \rho_i$ $(i=1,2)$. Under this assumption, the ratio of marginal social costs (i.e., the slope of the production possibility curve) is equated to the ratio of marginal private costs (i.e., the price ratio). The producer equilibrium conditions of (3.24) now result in both production efficiency and product mix efficiency.

4. Market Equilibrium

Urbanization diseconomies introduced by the consumption sector, and localization economies introduced by the production sector can now be combined to show the effect of the net urbanization economies on the city economy in equilibrium. Our single–city model can be summarized into three sets of equations. From (3.12), the consumer equilibrium conditions,

$$p_i X_i = w\theta_i N\left(1 - \frac{gN}{2\pi} + \frac{rK}{wN}\right) \quad (i=1,2), \quad (3.12)$$

from (3.24), the producer equilibrium conditions,

$$p_i Y_i = \frac{wL_i}{\alpha_i} = \frac{rK_i}{1-\alpha_i} \qquad (i=1,2), \qquad (3,24)$$

and the production possibility set, which consists of : the objective production functions ; and from (3.15), the city's resource endowment,

$$Y_i = \left(L_i^{\alpha_i} K_i^{1-\alpha_i}\right)^{\rho}, \quad \Sigma L_i = N\left(1 - \frac{gN}{2\pi}\right),$$
$$\Sigma K_i = K. \qquad (3.27)$$

The model is now colsed by the condition, $X_i = Y_i$. Appropriate substitutions will determine the equilibrium factor price ratio, and the allocation of factors :

$$\frac{r}{w} = \frac{\Sigma(1-\alpha_i)\,\theta_i}{\Sigma\,\alpha_i\,\theta_i}\left(1 - \frac{gN}{2\pi}\right)\left(\frac{K}{N}\right)^{-1},$$

$$L_i = \frac{\alpha_i\theta_i}{\Sigma\alpha_i\theta_i} N\left(1 - \frac{gN}{2\pi}\right), \; K_i = \frac{(1-\alpha_i)\theta_i}{\Sigma(1-\alpha_i)\theta_i} K,$$
$$(i=1,2). \qquad (3.28)$$

The output can now be obtained by substituting (3.28) into the objective production functions (3.27) :

$$Y_i = X_i = \left[\left(\frac{\alpha_i}{\Sigma\alpha_i\theta_i}\right)^{\alpha_i}\left(\frac{1-\alpha_i}{\Sigma(1-\alpha_i)\theta_i}\right)^{1-\alpha_i}\left(\frac{K}{N}\right)^{1-\alpha_i}\right.$$
$$\left.\theta_i N\left(1 - \frac{gN}{2\pi}\right)^{\alpha_i}\right]^{\rho} \quad (i=1,2). \quad (3.29)$$

The prices sre now derived by substituting (3.28) and (3.29) into the demand functions (3.12) :

$$\frac{p_i}{w} = c_i\left(\theta_i N\right)^{1-\rho}\left(1 - \frac{gN}{2\pi}\right)^{1-\rho\alpha_i}$$
$$(i=1,2), \qquad (3.30)$$

where c_i = constant term of parameters. Due to the homogeneous nature of our model, only the ratio pi/w, that is, the real price of good i in terms of labor, is determinate.

We now show that the equilibrium is both unique and stable by means of the supply and demand curves. Combining (3.24), (3.27), and (3.28), we obtain the supply curves,

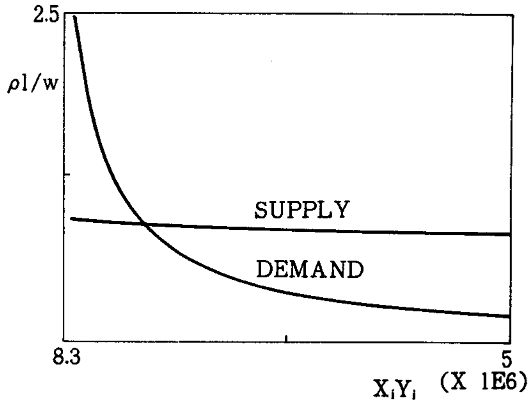
$$\frac{p_i}{w}=\alpha_i^{-\alpha_i}\left((1-\alpha_i)\frac{\Sigma\alpha_i\theta_i}{\Sigma(1-\alpha_i)\theta_i}\,\frac{K}{N}\right)^{\alpha_i-1}\left(1-\frac{gN}{2\pi}\right)^{1-\alpha_i}Y_i^{(1-\rho)/\rho}\qquad (i=1,2).\qquad (3{,}21)$$

and combining (3.12) and (3.28), we obtain the demand curves,

$$\frac{p_i}{w}=\frac{\theta_i}{\Sigma\alpha_i\theta_i}N\left(1-\frac{gN}{2\pi}\right)\frac{1}{X_i}\quad (i=1,2).\quad (3.32)$$

Both curves and the solution are depicted in fig. 3—1 for industry 1. The supply curve has constant elasticity $\rho/(1-\rho)$, and the demand curve has constant elasticity -1. Since $\rho/(1-\rho)<-1$ with $\rho>1$, the supply curve is always flatter than the demand curve, and they meet exactly once. Therefore, the equilibrium is always unique and stable.

In order to check rough orders of magnitude for the equilibrium, some tentative values are assigned to the parameteers as in fig, 3—1. First, we assume that the upper limit of N of (3.16) to be 10^7 so that $g/2\pi=0.25\times10^{-7}$. For a city of size $N=1\times10^6$, for example, this means that 2.5 percent of the total time is used for commuting, and $L=0.975\times10^6$. The critical value, however, is to come from the degree of homogeneity ρ. In a rare attempt to measure the extent of localization economies in the US SMSA data of the two-digit SIC manufacturing industries, Shefer [1973] reports the median values of ρ to



$\alpha_1=0.7\quad \theta_1=0.5\quad \rho=1.04\quad N=1\times10^6$

$K=2\times10^6\quad \alpha_2=0.3\quad \theta_2=0.5g/2_z=0{,}25\times10^{-7}$

In equilibrium, $X_1=Y_1=1.122\times10^6$, $p_1/w=0.896$.

Fig. 3—1 : Supply and Demand Curves (i=1).

be roughly between 1.03 and 1.05. Other US studies based on urbanization economies, however, report slightly different values (Sveikauskas[1975], Segal[1976], Moomaw [1981]). We initially choose $\rho=1.04$.

The values of labor elasticities in the subjective production function $\alpha_1=0.7$, $\alpha_2=0.3$ were deliberately chosen to allow the two industries to have markedly different production technologies, which along with "mild" scale economies, generally tend to retain the familiar concavity in the production possibility curve despite scale economies (Kemp[1964], Melvin[1969]). With $\theta_1=\theta_2=0.5$, and $K/N=2$, we get for the first industry, for example, $L_1=6.825\times10^5$, $K_1=6\times10^5$, $p_1/w=0.869$, and $X_1=Y_1=1.122\times10^6$. The output is remarkably large, because under constant returns to scale, i. e. $\rho=1$, it would be only 0.657×10^6. Despite our allowance for multiple industries in the city, the existence of only two industries in a big city seems to cause considerable specialization by each industry.

Equation (3.30) shows that the equilibrium price is a function of K/N, N, and a set of production, consumption, and transportation parameters. When K/N increases at given N, the price curve shifts down. This is a normal factor proportion effect accentuated here with localization economies. If we isolate the effect of K/N, the effect of city size N on the price can be identified. Taking the derivative of (3.30) with respect to N, while holding K/N fixed, we arrive at the following elasticity of price with respect to city size,

$$\frac{d(p_i/w)}{dN}\ \frac{N}{(p_i/w)}=\frac{\rho(1-\alpha_i)\frac{gN}{2\pi}}{1-\frac{gN}{2\pi}}<0. \qquad (3.32)$$

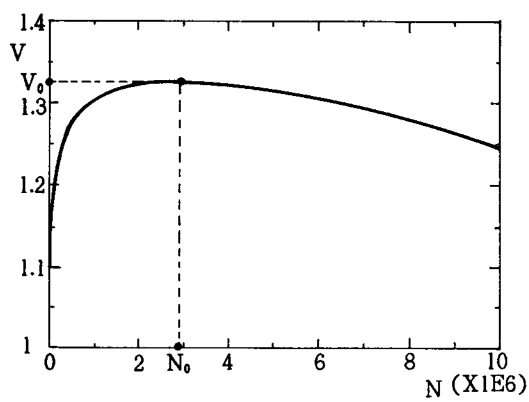
As expected, the equilibrium price becomes a monotonically decreasing function of N. With equi-proportional increases in K and N, the city will experience unequivocal decreases in prices. Higher degrees of homogeneity ρ and greater capital elasticities in the subjective production function $(1-\alpha_i)$ will increase the above elasticity in absolute terms, and accentuate the relationship.

We are now ready to determine the equilibrium utility level in the city. Substituting the factor price ratio (3.28) and the prices of goods (3.30) back into the indirect utility function (3.13), we get

$$v=c_2c_3N^{\rho-1}\left(1-\frac{gN}{2\pi}\right)^{\rho\Sigma\alpha_i\theta_i}, \qquad (3,33)$$

where c_2, c_3=constant terms of parameters. Unlike the price case of (3.28) in which urbanization economies are largely reflected, the effect of N on the utility level is not always in the form of reduced labor with increased city size, which not only decreases the total resource base of the city, but also in creases the capital-labor ratio. Since decreasing returns to factor proportions are assumed, increase in N eventually leads to diminished industry output. Thus, the utility level of (3.33) also represents the equlilbrium net urbanization economies for the city. This relationship is presented in Fig. 3-2. With increase in N, the utility level initially goes up rather rapidly, attains a maximum, and eventually goes down in a less rapid, asymmetrical manner.

It is to be noted that each point on the curve of Fig. 3-2 indicates both an equilibrium and an optimum, not just in the private sense but, on the account of $\rho=\rho_i$,



$\alpha_1=0.7$ $\theta_1=0.5$ $\rho=1.04$

$\alpha_2=0.3$ $\theta_2=0.5$ K/N=2 $g/2\pi=0.25\times10^{-7}$

$N_0=2.86\times10^6$ $V_0=1.324$

Fig. 3-2 : Utility Level with City Size.

in the social sense as well. The curve shows, for the city in autarky, the highest level of welfare attainable under the production possibility set, which is determined by the given city size and other internal characteristics of the city. On the other hand, the "optimum city size", which corresponds to N_0 of Fig. 3-2 relates to the global maximum of the curve. This particular size N_0 and the resulting production possibility set assure that the city output and welfare are maximized in real terms on a per capita basis. It is found by $dv/dN=0$ from (3.33), while holding K/N fixed, and turns out to be

$$N_0=\frac{2\pi}{g}\,\frac{\rho-1}{(\rho-1)+\rho\Sigma\alpha_i\theta_i}\,. \qquad (3.34)$$

For the illustrative parametric values employed in Fig. 3-2, $N_o=2.86\times10^6$, to arrive at $V_o=1.324$. This size N_o is below the practical upper limit of N, assumed to be 10^7, and therefore remains as an attainable target.

We have now completed modelling of our single-city with multiple industries, and proposed that urbanization economies and diseconomies be measured in utility terms. Unlike other single-city models,our model indicates that in equilibrium, the criteria of efficiency are met, i. e., efficiency in consumption, production, and output mix.

5. Migration and City Size Distribution

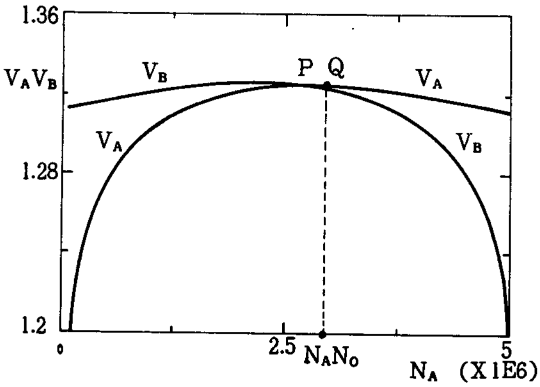
Like the models of the supply-oriented approach in general, our single-city model was concerned with one city in isolation, and not with its relationship with other cities. In this section, we establish a limited degree of interdependence between cities and derive a resulting city size distributiion by allowing free factor movements but not trade between cities. For this, we assume that all cities share identical production, utility, and commuting cost functions. We further suppose that people migrate costlessly from city to city to maximize utility, that each migrant in the nation owns capital equally and takes with him his share, and that capital rentals as well as wages are spent in the city where people work and live. Notice that the movement of people does not change the ratio of capital to population in either the source city or the host city. The utilization of resources, however, will generally differ between cities of different sizes.

Now suppose initially that two cities A and B are engaged in the exchange of factors by following the rules listed above. With people (along with their capital) moving in response to utility differences, an equilibrium is reached when a common utility level is achieved in the two cities. Equation (3.33) is directly applicable, and the equilibrium conditions are

$$V_A=V_B, \qquad N_A+N_B=\overline{N}, \qquad (3.35)$$

Where : V_A, V_B=utility levels in cities A and B respectively,
C_2,C_3=constant terms of parameters,
N_A,N_B=populations of cities A and B respectively,
$\overline{N}$=population in the nation.

The shapes of the utility curves and the solution for (3.35) generally depend on the national population parameter N as it



$\alpha_1=0.7$ $\theta_1=0.5$ $\rho=1.04$

$\alpha_2=0.3$ $\theta_2=0.5$ $\bar{K}/\bar{N}=2$ $g/2\pi=0.25\times10^{-7}$

$N_0=2.86\times10^6$, $\bar{N}=5\times10^6$, $N_A=N_B=2.5\times10^6$

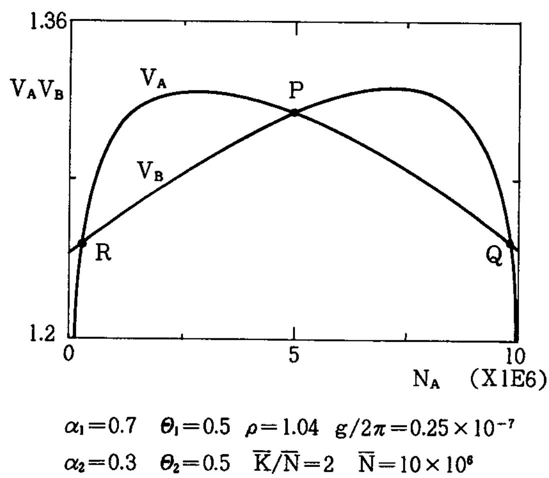
Fig. 3-3 : Equilibrium City Sizes($\bar{N}\leqq2N_o$).

relates to the optimum city size N_o of the single-city model. When N is sufficiently small so that $\bar{N}\leqq2N_o$ a unique equilibrium is reached at point P in Fig. 3-3, in which the horizontal axis

is fixed of length $\bar{N}$, and measures N_A from left to right and N_B the other way around. Due to the symmetry between the curves V_A and V_B, however, we should suspect that this equilibrium is generally unstable, because slight migration from city B to A due to an exogenous shock will increase V_A as well as decrease V_B. The gap between V_A and V_B then becomes cumulative as the migration progresses. The migration will continue beyond Q at which V_A attains its maximum and N_A becomes N_o.

It is also clear that even when $N_A=N_B=N_o$, the equilibrium is unstable due to the same disequilibrating migratory flow. Of course, when $\bar{N}$ is even smaller, for exemple $\bar{N}\leqq\bar{N}_o$, there is no justificaton for the existence of the two citiis, and the migration will proceed until city B is totally vacated so that $N_A=\bar{N}<N_o$. However, this is an extreme case which can hardly be imaginable in reality. To sum up, it is therefore clear that a unique and stable equilibrium in the city size distribution when $\bar{N}\leqq2N_o$ is not possible. A numerical example is given in Fig. 3-3 for the unstable equilibrium at P.

When $\bar{N}>2N_o$, multiple equilibria appear to be reached at three different sets of city sizes such as points P, Q, and R in Fig. 3-4, it represents a highly skewed city size distribution at $N_A=9.8\times10^6$, and $N_B=0.2\times10^6$. However, the equilibrium at Q is most unstable, because any movement toward another equilibrium point P unambiguously improves the utility levels for both cities. Migration from city A to B, once started, will proceed until point P is reached at which $N_A=N_B>N_o$. The equil-



$\alpha_1=0.7$ $\theta_1=0.5$ $\rho=1.04$ $g/2\pi=0.25\times10^{-7}$

$\alpha_2=0.3$ $\theta_2=0.5$ $\bar{K}/\bar{N}=2$ $\bar{N}=10\times10^6$

At P, $N_A=N_B=5\times10^6$; at Q, $N_A=9.8\times10^6$, $N_B=0.2\times10^6$

Fig. 3-4 : Equilibrium City Sizes ($\bar{N}>2N_o$).

ibrium at P which is on the falling part of both V_A and V_B, however, must be a stable one. A slight deviation from P toward Q will be met by counter−migration from city A to B, thereby restoring the equilibrium. When $\bar{N}>2N_0$, therefore, a stable equilibrium is reached at P where city sizes are identical but larger than N_0 ($N_A=N_B>N_0$). Since the common ntility level in the equilibrium falls as $\bar{N}$ increases, however, there must be a practical limit in this equilibrium. It would be unreasonable to operate under the confines of two cities in the nation when $\bar{N}$ is much greater than $2N_0$.

If we extend the above further to a more general situation in which $\bar{N}$ is much larger than $2N_0$ and the existence of more than two cities are considered, however, we suspect that the stable equilibrium city size will be N_0 for all cities in the nation. As Henderson[1977] argues, this is because with $\bar{N}$ being sufficiently large, the divisibility problem due to lumpiness can be avoided in replication of cities with size N_0. For example as in Fig. 3−4, suppose there are many cities with sizes slightly larger and smaller than N_0. With many cities in existence, the sum of any "excess" populations beyond N_0 from the larger cities could be added to the smaller cities, thereby helping to achieve N_0 among all cities. Thus, all cities eventually converge N_0 in Size.

It is argued that this rather "trivial" distribution of city sizes in equilibrium is an inherent phenomenon in the models of the supply−oriented approach. The lack of opportunity to trade goods forces the city to produce locally what it consumes under the equilibrium conditions that reflect the internal characteristics only. With no consideration given to other cities in both production and consumption, all cities tend to be "standardized", and are required to converge on the ideal city size N_0 for equilibrium. This is in a sharp contrast with the reality in general, and the trade model in particular, in which cities of continuously different sizes can coexist and benefit from each other with trade albeit in a suboptimal manner.

Ⅳ. A TRADE MODEL WITH SCALE ECONOMIES

1. Introduction

The purpose of this chapter is to allow trade between the two cities of our single−city, and to review the resulting city size distribution in the nation. Under the assumption that the nation consists of the two cities, the inter−city terms of trade ($p=p_2/p_1$) becomes endogenous, and is to be determind in equilibrium. While housing consumption, commuting, and the labor supply will still depend on the city's internal characteristics, the production and consumption of traded goods, and the level of welfare will also depend on those of the trading partner.

This will be carried out in two steps. First, we will develop a balanced free−trade model with such requirement as free trade, and balance of payments. In this step, migration is not allowed, and city sizes are fixed. Under urbanization economies and diseconomies, although commodity prices will be the same in equilibrium, utility levels will in general differ between the two cities. The next step will be to allow factor movements to arrive at a common level of utility. Given the national population, this will allow us to derive a system of cities of generally different sizes. Despite scale economies, the model wili operate much like the standard trade model of factor proportions. This is the result of our assumption in which "mild" scale economies for both industries give rise to incomplete specialization for the city.

2. Production under "Mild" Scale Economies

When the degress of scale economies are sufficiently large, the marginal social opportunity cost that measures the slope of the production possibility curve of the city decreases, and the production possiblility curve has a convex shape. Scale then becomes the dominant basis for trade, and the usual stable equilibrium results in complete specialization by each city in one yet indeterminate good. This situation, however, seems hardly plausible in the case of exteral economies, because it would be naive to assume that the entrepreneurs behave competitively despite such high degrees of scale economies.

On the other hand, "mild" scale economies combined with a large difference in the factor intensities yield the concave production possibility curve for the most part except for the extreme points close enough to both axes. Now comparative advantage arising from different factor proportions between cities of different sizes becomes the basis for trade. Because the marginal social opportunity cost is rising, the pattern of trade will generally show more modest specialization. In short, imcomplete specialization is more likely under moderate degress of scale economies. In light of urban industrial diversity in reality, we assume that the city is subject to the concave production possibility curve at least for its relevant part.

The assumption of the concave production possibility curve requires that $d^2Y_1/dY_2^2=d(-p)/dY_2<0$, where p is the price of the second good in terms of the first $(p=p_2/p_1)$. To probe it further, we first express endogenus variables in terms of q defined below. Form (3.29), the factor price ratio is equal to the marginal rates of susubstitution in production, which, in turn, depend on the factor ratios alone, regardless of the scale of output :

$$\frac{r}{w}=\frac{1-\alpha_i}{\alpha_i}\left(\frac{K_i}{L_i}\right)^{-1}. \qquad (4.1)$$

Along with the full employment conditions from (3.27), and letting

$$q=\frac{1-\dfrac{gN}{2\pi}+\dfrac{rK}{wN}}{1-\dfrac{gN}{2\pi}}, \qquad (4.2)$$

where : $1<\dfrac{1}{\alpha_1}\leq q\leq\dfrac{1}{\alpha_2}$, $0<\alpha_2<\alpha_1<1$,

we derive the following :

$$L_1=\frac{\alpha_1(1-q\alpha_2)}{\alpha_1-\alpha_2}N\left(1-\frac{gN}{2\pi}\right),$$
$$L_2=\frac{\alpha_2(q\alpha_1-1)}{\alpha_1-\alpha_2}N\left(1-\frac{gN}{2\pi}\right). \qquad (4.3)$$
$$\frac{K_i}{L_i}=\frac{1-\alpha_i}{\alpha_i}\frac{K}{N}(q-1)^{-1}\left(1-\frac{gN}{2\pi}\right)^{-1}.$$

Finally Substituting (4.3) into (3.24) and (3.27), the remaining endogenous variables are expressed in q :

$$Y_1=\left[\alpha_1^{\alpha_1}(1-\alpha_1)^{1-\alpha_1}\left(\frac{K}{N}\right)^{1-\alpha_1}\frac{N}{\alpha_1-\alpha_2}\right.$$

$$\left.1-\left(\frac{gN}{2\pi}\right)^{\alpha_1}(1-q\alpha_2)\ (q-1)^{\alpha_1-1}\right]^{\rho},$$

$$Y_2=\left[\alpha_2^{\alpha_2}(1-\alpha_2)^{1-\alpha_2}\left(\frac{K}{N}\right)^{1-\alpha_2}\frac{N}{\alpha_1-\alpha_2}\right.$$

$$\left.\left(1-\frac{gN}{2\pi}\right)(q\alpha_1-1)\ (q-1)^{\alpha_2-1}\right]^{\rho},$$

$$p=c_1\left(\frac{K}{N}\right)^{-\rho(\alpha_1-\alpha_2)} \tag{4.4}$$

$$\left[\left(1-\frac{gN}{2\pi}\right)(q-1)\right]^{\rho(\alpha_1-\alpha_2)}\left(\frac{1-q\alpha_2}{q\alpha_1-1}\right)^{\rho-1}, \tag{4.5}$$

where : c_1 = a constant term of parameters. As (4.2) indicates, when the city is completely specialized in the labor-intensive first good (or capital-intensive second good), q becomes $1/\alpha_1$ (or $1/\alpha_2$). When it is incompletely specialized, q takes an intermediate value.

Differentiating the endogenous variables with respect to q, we derive the following elasticities :

$$\frac{dY_1}{dq}\frac{q}{Y_1}=-\rho q\left(\frac{\alpha_2}{1-q\alpha_2}+\frac{1-\alpha_1}{q-1}\right)<0,$$

$$\frac{dY_2}{dq}\frac{q}{Y_2}=\rho q\left(\frac{\alpha_1}{q\alpha_2-1}-\frac{1-\alpha_2}{q-1}\right) \tag{4.6}$$

$$=\rho q\left(\frac{1-\alpha_1+\alpha_2(q\alpha_1-1)}{(q\alpha_1-1)\ (q-1)}\right)>0,$$

$$\frac{dp}{dq}\frac{q}{Y_2}=q(\alpha_1-\alpha_2)$$

$$\left(\frac{\rho}{q-1}-\frac{\rho-1}{(q\alpha_1-1)(1-q\alpha_2)}\right)>0 \tag{4.7}$$

Since $d(-p)/dY_2=-(dp/dq)\ (dq/dY_2)$, and $dY_2/dq>0$ from (4.6), if $dp/dq>0$, which from (4-7) is very likely when $(\rho-1)$ is sufficiently small and close to zero, then $d(-p)/dY_2<0$ and the production possibility curve becomes concave. Under constant returns to scale, i. e., $\rho-1=0$, the price ratio (4.5) depends solely on the factor price ratio that determines the factor proportions via (4.3). Under increasing returns to scale, however, it also depends on the relative degree of specialization (measured by the term with the exponent $(\rho-1)$. With small $(\rho-1)$, the factor proportions effect will tend to dominate over the "specialization effect".

3. A Balanced Free-Trade Model

We now consider balanced free-trade between the two cities A and B of fixed sizes comprise the nation. We assume that city A is larger than B, i. e., $N_A>N_B$. One of the general equilibrium conditions is that terms of trade be equal. Rearraging (4.5), and assigning q_A and q_B for q to cities A and B respectively, we obtain

$$p_A=C_4\left[\left(1-\frac{gN_A}{2\pi}\right)(q_A-1)\right]^{\rho(\alpha_1-\alpha_2)}\left(\frac{1-q_A\alpha_2}{q_A\alpha_1-1}\right)^{\rho-1},$$

$$p_B=C_4\left[\left(1-\frac{gN_B}{2\pi}\right)(q_B-1)\right]^{\rho(\alpha_1-\alpha_2)}\left(\frac{1-q_B\alpha_2}{q_B\alpha_1-1}\right)^{\rho-1}, \tag{4.8}$$

$$p_A=p_B,$$

where : p_A, p_B = price ratios of good 1 to good 2 in A and B,

c_4 = a constant term of parameters,

$$\frac{r_A}{w_A}=(q_A-1)\left(1-\frac{gN_A}{2\pi}\right)\left(\frac{K_A}{N_A}\right)^{-1},$$

$$\frac{r_B}{w_B}=(q_B-1)\left(1-\frac{gN_B}{2\pi}\right)\left(\frac{K_B}{N_B}\right)^{-1}, \tag{4.9}$$

where : r_A, r_B, w_A, w_B=capital rentals, wage rates of A and B.

Equation (4.8) shows that the price ratios are increasing functions of q when $dp/dq>0$. It is clear that if city sizes are indentical, $N_A=N_B$, then the autarky price ratios are the same, i. e., $p_A=p_B$ at $q_A=q_B=1/\Sigma\alpha_i\theta_i$, and there is no basis for trade. Only when city sizes are different, the autarky price ratios become different, and they will set the limits for the equilibrium terms of trade. In equilibrium, therefore, q_A must be greater and q_B must be less than $1/\Sigma\alpha_i\theta_i$. This, along with (4.2), gives the ranges of q_A and q_B :

$$1<\frac{1}{\alpha_1}\leq q_B\leq\frac{1}{\Sigma\alpha_i\theta_i}\leq q_A\leq\frac{1}{\alpha_2}. \qquad (4.10)$$

Because of (4.10), imcomplete specialization by both cities is practically assured. This is because one city's degree of specialization, in the sense of the shift from autarky to the equilibrium production, is to be limited by that of the other. In short, specialization is limited by the extent of the market.

In addition to (4.8), equilibrium requires that trade be balanced. Because inter-city investment and transport costs are not considered, this requires that the national demand for any good must be equal to the national supply. Already knowing the pattern of trade, we define the net excess supply of the second good by city A, E_{2A}, as the difference between the local supply, Y_{2A}, and the local demand, X_{2A}. From (3.12), (3.24) (4.3),

$$E_{2A}=Y_{2A}-X_{2A}=\frac{w_A}{p_{2A}}\ \frac{q_A\Sigma\alpha_i\theta_i-1}{\alpha_1-\alpha_2}$$

$$N_A\left(1-\frac{gN_A}{2\pi}\right)\geq 0,\ \text{or} \qquad (4.11)$$

$$E_{2A}=C_5\ N_A^{\rho}\left(1-\frac{gN_A}{2\pi}\right)^{\rho\alpha_2}(q_A-1)^{\rho(\alpha_2-1)}$$

$$(q_A\alpha_1-1)^{\rho-1}\ (q_A\Sigma\alpha_i\theta_i-1),$$

where c_5=a constant term of parameter.

Likewise, we define the net excess demand for the second good by city B, E_{2B}, and express it in terms of q_B.

$$E_{2B}=X_{2B}-Y_{2B}=\frac{w_B}{p_{2B}}\ \frac{1-q_B\Sigma\alpha_i\theta_i}{\alpha_1-\alpha_2}$$

$$N_B\left(1-\frac{gN_B}{2\pi}\right)\geq 0, \qquad (4.12)$$

$$E_{2B}=C_5\ N_B^{\rho}\left(1-\frac{gN_B}{2\pi}\right)^{\rho\alpha^2}(q_B-1)^{\rho(\alpha_2-1)}$$

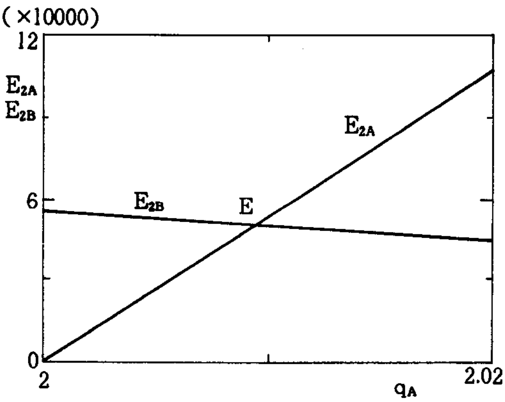
$$(q_B\alpha_1-1)^{\rho-1}\ (1-q_B\Sigma\alpha_i\theta_i).$$

Hence, the balance of trade :

$$E_{2A}=E_{2B}. \qquad (4.13)$$

The solution in equilibrium can be obtained by solving the two sets of equations regarding the equalization of the price ratios (4.8), and the balance of trade consisting of (4.11), (4.12) and (4.13) simultaneously for the two unknowns q_A and q_B. Due to difficulty in solving them algebraically we will approach them by numerical iterative methods. The solution is presented in Fig. 4-1 and 4-2 for the parametric values employed therein. A unique equilibrium is therefore obtained at E that clears the markets in both cities for the second good, and on the account of Walras' Law, for the first good as well. By the virtue of the slopes of the E_{2A} and E_{2B} curves ($dE_{2A}/dp>0$, $dE_{2B}/dp<0$ and p=

$p_A = p_B$), it is clear that the equilibrium is also stable (i. e., $dE_{2B}/dp - dE_{2A}/dp < 0$). The solutions are : $q_A = 2.0093$, $q_B = 1.9090$, $p_A = p_B = 0.6137$, $E_{2A} = E_{2A} = E_{2B} = 5.024 \times 10^4$. As was expected, the deviation of q_B from the autarky point is far greater than that of q_A (0.0901 vs. 0.0093). Thus, despite the significant shift away from the autarky situation in city B, the situation in city A has barely been changed, thereby limiting the overall trade volume.



$\alpha_1 = 0.8$ $\alpha_2 = 0.2$ $\theta_1 = \theta_2 = 0.5$ p $= 1.03$

$N_A = 4 \times 10^6$ $N_B = 0.4 \times 10^6$

$\frac{K_A}{N_A} = \frac{K_B}{N_B} = 1$ $\frac{g}{2\pi} = 0.25 \times 10^{-7}$

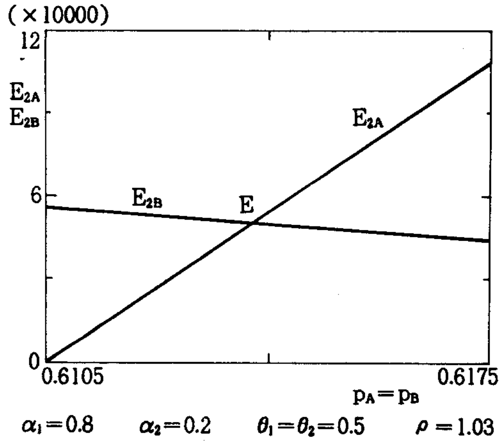
In equilibrium at E : $q_A = 2.0093$

$q_B = 1.90991$

$p_A = p_B = 0.61372$ $E_{2A} = E_{2B} = 5.024 \times 10^4$

〈그림 4－1〉 Numerical Solution of $p_A = p_B$, q_A vs. E_2

We now turn to the gains from trade in each city by looking at the improvement in the level of utility. For the representative city, the indirect utility function (3. 13) is denoted as a function of q by the use of (4.2),



$\alpha_1 = 0.8$ $\alpha_2 = 0.2$ $\theta_1 = \theta_2 = 0.5$ $\rho = 1.03$

$N_A = 4 \times 10^6$ $N_B = 0.4 \times 10^6$

$\frac{K_A}{N_A} = \frac{K_B}{N_B} = 2$ $\frac{g}{2\pi} = 0.25 \times 10^{-7}$

In equilibrium at E : $q_A = 2.0093$

$q_B = 1.90991$

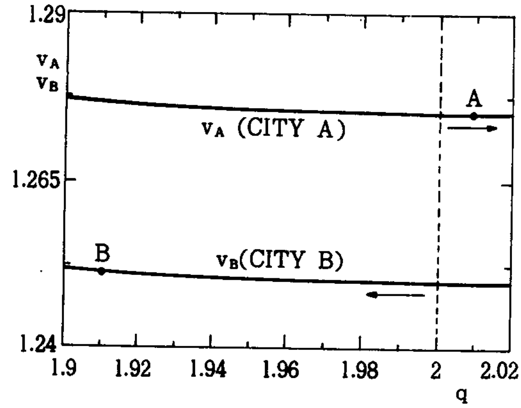
$p_A = p_B = 0.61372$ $E_{2A} = E_{2B} = 5.024 \times 10^4$

〈그림 4－2〉 Numerical Solution of $E_{2A} = E_{2B}$. p vs. E_2

$$v = c_6 N^{\rho - 1}\left(1 - \frac{gN}{2\pi}\right)^{\rho \Sigma \alpha \theta_i} q(q-1)^{-\rho \Sigma (1-\alpha_i)\theta_i} \left[(1 - q\alpha_2)^{\theta_1} (q\alpha_1 - 1)^{\theta_2} h\right]^{\rho - 1} \qquad (4.14)$$

where : c_6 = a constant term.

The above relationship along with the utility levels of both cities in both autarky and equilibrium is illustrated in Fig. 4－3 in which the arrow marks indicated the oppoisite directions of movement by the cities from autarky to equilibrium. Because the equilibrium point B for city B is located considerably farther away from the autarky point than its counterpart A for city A, the improvement in welfare is accordingly greater for the smaller city. While both cities gain from trade, this suggests that



$\alpha_1=0.8$ $\alpha_2=0.2$ $\theta_1=\theta_2=0.5$ $\rho=1.03$

$N_A=4\times10^6$ $N_B=0.4\times10^6$

$\frac{K_A}{N_A}=\frac{K_B}{N_B}=2$ $\frac{g}{2\pi}=0.25\times10^{-7}$

In equilibrium at A and B : $q_A=2.0093$
$q_B=1.9099$

$v_A=1.27553$ $v_B=1.25156$

Fig. 4−3 Changes in Utility Levels, Before and After Trade.

not only is the smaller city more dependent on trade, the gains from trade is also greater for the smaller city.

4. Migration and Balanced Free−Trade

The last section showed that while the utility levels of both cities were improved after trade, they were not equalized, thereby leaving a tendency for factors to move. In this secton, we incorporate factor movements into the model, so that city sizes become endogenous, and are to be determined in equilibrium given the exogenous variable $\overline{N}$, the national populaton. Because both trade and factor movements are simultaneously considered under the general equilibrium setting, this model represents a more direct and realistic rendition of the interaction between cities.

For this, an additional behavioral assumption is in order. The location decision of the capital owner/laborer becomes multidimensional. In his costless move from to city to maximize utility, he has to consider not only the marginal product of labor, but also the marginal product of capital and the net housing costs (after the rebate from the land bank), all specific to each city. Because captial always moves with the laborer under our assumption, we will use the term "migration" to denote the combined movements of factors.

City sizes being endogenous now, we define a new variable z as population share of the larger city A out of the national population $\overline{N}$ in equilibrium. With two cities in the nation, we obtain that

$$z=N_A/\overline{N}, \qquad 1-z=N_B/\overline{N}, \qquad 0<1-z\leq 1/2\leq z<1. \tag{4.15}$$

We can determine the equilibrium of our migration and trade model by solving simultaneously three equations regarding : the equalization of the product price ratios, $p_A=p_B$; the balance of trade, $E_{2A}=E_{2B}$; and the equalization of the utility levels, $v_A=v_B$. The last equation, however, needs further discussion.

For both cities, the indirect utility function (3.13) can be rewritten by the use of (4.2), (4.15) and (4.11) as

$$V_A=\theta_1^{\theta_1}\theta_2^{\theta_2}(\alpha_1-\alpha_2)(\overline{N})^{-1} p_A^{\theta_1}E_{2A}\frac{q_A}{z(q_A\Sigma\alpha_i\theta_i-1)}, \tag{4.16}$$

$$V_B = \theta_1^{\theta_1}\theta_2^{\theta_2}(\alpha_1-\alpha_2)\ (\overline{N})^{-1}\ p_B^{\theta_1} E_{2B}\ \frac{q_B}{(1-z)\ (1-q_B\Sigma\alpha_i\theta_i)}$$

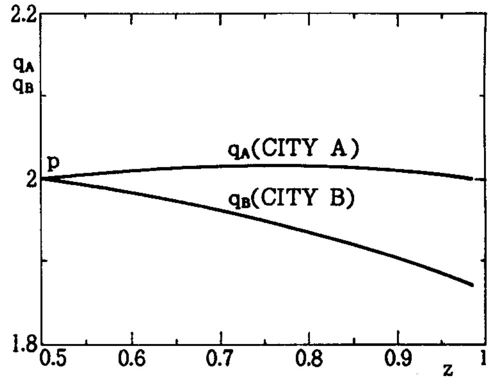
in which, provided the price ratios are equalized ($p_A=p_B$) and the trade is balanced ($E_{2A}=E_{2B}$), the utility levels are equalized only if

$$\frac{z}{q_A}+\frac{1-z}{q_B}=\Sigma\alpha_i\theta_i,\ \text{or},\ q_B=\frac{q_A(1-z)}{q_A\Sigma\alpha_i\theta_i-z}. \qquad (4.17)$$

which further limits the deviations of q_A and q_B from their autarky value, $1/\Sigma\alpha_i\theta_i$.

By utilizing (4.17), we can thus reduce one variable and express the relevent equations in terms of two unknowns. Equation $p_A=p_B$ was numerically solved first, and its solutions were applied to equation $E_{2A}=E_{2B}$. Both are presented in Fig. 4−4 and 4−5 respectively. Accroding to Fig. 4−5 or 4−6 the latter of which shows the relationship between the utlity levels of the two cities vs. z, we have two equilibrium points : whereas P refers to the autarky equilibrium with identical city sizes (z=0.5), Q refers to the trade equilibrium with different city sizes (z=0.75793).

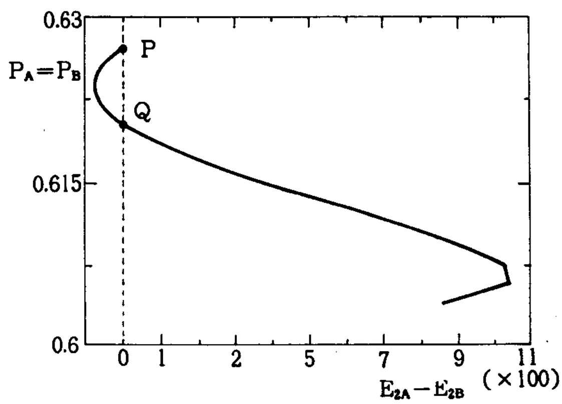
Are they stable? Because the "trivial" equilibrium at P involves no trade and is determined entirely by migration, its stability is determined by the shapes of the utility curves of the two cities as they are related to their respective city size. As was noted in the last chapter, the equilibrium at P is stable only if the equilibrium city sizes are greater than the optimum city size of the signle−city model N_0 which at the current parametric values is 2.201 $\times 10^6$. At P Fig. 4−8, for example, $N_A = N_B = 0.5\overline{N} = 2.4\times10^6 > N_o$ and it is thus stable.



$\alpha_1=0.8 \quad \alpha_2=0.2 \quad \theta_1=\theta_2=0.5 \quad \rho=1.03$

$\overline{N}=4.8\times10^6 \quad \overline{K}/\overline{N}=2 \quad \frac{g}{2\pi}=0.25\times10^{-7}$

〈그림 4−4〉 Numerical Solution of $V_A=V_B$ and $P_A=P_B$, z vs. q.



$\alpha_1=0.8 \quad \alpha_2=0.2 \quad \theta_1=\theta_2=0.5 \quad p=1.03$

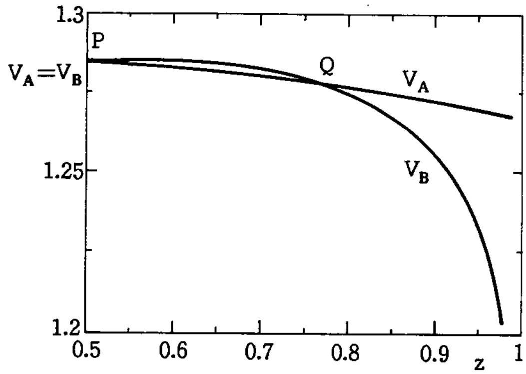
$\overline{N}=4.8\times10^6 \quad \overline{K}/\overline{N}=2 \quad \frac{g}{2\pi}=0.25\times10^{-7}$

At P, $E_{2A}=E_{2B}=0 \quad P_A=P_B=0.62713$

At Q, $E_{2A}=E_{2B}=8.5373\times10^4$

$P_A=P_B=0.62037$

〈그림 4−5〉 Numerical Solution of $V_A=V_B$ and $E_{2A}=E_{2B}$, $E_{2A}-E_{2B}$ vs. p.



$\alpha_1=0.8$ $\alpha_2=0.2$ $\theta_1=\theta_2=0.5$ $p=1.03$
$\overline{N}=4.8\times10^6$ $\overline{K}/\overline{N}=2$ $\frac{g}{2\pi}=0.25\times10^{-7}$
At P, $z=0.5$ $V_A=V_B=1.28449$
At Q, $z=0.75793$ $V_A=V_B=1.27843$

〈그림 4−6〉 Equilibrium Utility Levels Vs City Size Distribution

The equilibrium at Q, on the other hand, concerns both trade and migration, and its stability test requires the examination of both Fig. 4−5 and 4−6. There, $d(E_{2A}-E_{2B})/dp$ is negative in the neighborhood of Q, so it is stable in terms of trade under the Marshallian criterion. However, it is unstable in terms of migration, because, according to Fig. 4−6 any migratory disturbance from Q will be cumulative in either direction, and set a stage for further disequilibrating migration away from the equilibrium. It is to be noted that whereas a more even city size distribution than the one at Q is welfare−improving for both cities, a more skewed distribution toward a greater primacy of the larger city is welfare−deteriorating for both cities.

While the equilibrium at P is meaningful and stable only when $\overline{N}>2N_0$ within the two−city framework, the equilibrium at Q cannot be obtained otherwise. This is because, as long as trade occurs, the balance of trade or equalization of the utility levels cannot be achieved unless $N_A>N_0$ and N_B N_0. Moreover as $\overline{N}$ increases, accommodation of the increased national population within the two cities requires that z must increase as well, leading toward a greater primacy. Because the utility levels at

P are always higher than those at Q at any relevant N, it is tempting to denote the point O (where $N_A=N_B=N_o$) as the "optimum national population".

V. POLICY IMPLICATIONS

Our findings can be summarized as follows.

(a). The optimum city size distribution in which potential level of welfare for all in the nation is maximized is characterized by the identical size N_0, the single-city optimum size, among all cities. A higher(lower) degree of homogeneity associated with localization economies ρ and/or a lower(higher) weighted average of the labor elasticities of the industry subjective production functions $\Sigma_{\alpha i}\theta_i$ will increase(decrease) N_0. It is also a stable equilibrium with many cities in existence in the economy. In the two-city framework, however, stability in equilibrium requires the common city size be larger than N_0, thereby causing a slight suboptimum compared to the theoretical optimum above. This optimum is achieved totally by migration of factors and absence of trade. This suggests that the common view that optimum city systems would probably be hierarchical in size such as the rank-size distribution(for example, Richardson[1981]) may be mistaken.

(b). For small and/or underdeveloped countries with a sufficiently low level of urbanization so that the total urban population is less than N_0, the optimum as well as equilibrium city size distribution is characterized by the existence of a single city in the nation due to cumulative migration. Because the level of welfare of the city is still on the rising part of the curve with respect to city size, the primary concern for such countries must be greater urbanization and industrialization before any consideration is taken into city size distribution.

If factor mobility is sufficiently low in such countries so that existing city sizes can be regarded as fixed, a trade equilibrium is feasible that offers gains from trade over autarky albeit at different welfare levels between cities. However, trade between small cities under scale economies is generally characterized by low trade volume and offers little improvement. Thus promoting greater factor mobility under the circumstances, thereby helping to achieve the single equilibrium city size, seems a lot more sensible approach than, for example, investment in inter-urban transportation systems that would presumably facilitate trade rather than factor mobility. Similarly, insitu development strategies aimed at the esisting small cities and towns such as the agropolitan development approach(Friedmann and Weaver[1979]) would appear ineffective.

(c). In the above situation (b), the existence of a single large city or its dominance over smaller cities in equilibrium must not be regarded as being of economic inefficiency. Although the city size distribution may appear highly primate, the dominant city is still small in absolute terms(i.e., less than N_0), and the equilibrium city size distribution represents close to the optimum. This suggests that relative primacy measures of city size distributions such as the ratio-based primacy index, or its variants such as the one by El-Shakhs[1972], or the statistically-determined Pareto coefficient (see Rosen and Resnick[1980]) all have inherent shortcomings in terms of associating economic efficiency with city size distributions. Clearly there is a need to complement them with some measures of absolute city sizes.

(d). The more general equilibrium city size distribution with different sizes is obtained by our migration and trade model, and it is characterized by the larger city being greater than and the smaller city less than N_0 within the two-city framework. It is required that the combined sizes be no less than $2N_0$. This situation is thus likely to emerge in countries which are larger in size and/or have a higher level of urbanization than those in (b). In equilibrium, trade is balanced and welfare levels are equalized between the two cities. This is in a sharp contrast with the common central-place-based hierarchical models, notably Beckmann and McPherson[1970] which is characterized by trade imbalance and a functional presupposition between the higher-order city size and its lower-order market areas.

It is to be noted that in order to arrive at this equilibrium, trade is a precondition ; otherwise, migration alone would result in a distribution of identical city sizes as in the above (a). It is in this sense that while the equilibrium is stable with respect to trade, it is unstable with respect to factor mobility. Although this equilibrium was shown in Fig. 4-6 to be inferior to the optimum city size distribution of (a), the almost universal occurrence of hierarchical city size distributions in reality leads us to suspect that factor mobility may not be as perfect as we have assumed. It may also be speculated that trade is so pervasive that migration occurs just enough to offset the post-trade welfare differentials between cities.

(e). The equilibrium city size distributionin (d) seems to closely reflect the current "primate" situation in many countries in that the very fact of hierarchical city sizes indicates a divergence from the optimum. Market outcomes under limited facor mobility clearly entail economic inefficiency, and a planned intervention that facilitates greater factor mobility and movement toward the optimum is consequently justified. In-deed, even, a small difference in size between the larger and smaller cities, although not primate at all, must be indicative of economic inefficiency.

Within the confines of two cities in the nation, the equilibrium situation gets worse and moves toward even greater primacy as the national urban population increases. However, the supposition is arbitrary and may exaggerate reality especially in view of the fact that many countries probably have national urban populations far greater than $2N_0$. Within the framework of many cities and sufficiently large urban populations, our model would instead predict an equilibrium in which multiples of pairs of different-sized trading cities exist. It is not difficult to envision that this new equilibrium may probably be closer to the optimum than the one under the two-city framework.

(f). It is well observed that with enough urbanization and/or development, urban primacy eventually falls and the importance of a city size distribution diminishes. And our last ovservation in (e) seems to support this. In connection with (d) and (e), however, for countries with not enough overall urbanization and/or development, the problem of urban primacy seems real and may indeed incur economic costs. Countries like Egypt, Mexico, and Brazil in which one or two large cities apper well beyond and other cities well below N_0 may belong to this group. Although further migration into the large cities may eventually slow down as the countries approach "lower level equilibrium points" in city size distributions, the utility levels in both large and small cities will deteriorate.

According to our model, a higher level equilibrium point under trade can be achieved only when the small cities grow larger in size but remain less than N_0. With the total national urban population fixed, this suggests that the smaller cities must be fewer in number as well. This suggestion seems in line with the paradigm of "concentrated decentralization" advocated by Rodwin[1961], and Alonso[1968], among others. Finally, the current near monopoly by the big cities of localized noneconomic functions, such as central government services and quality higher education, undoubtedly seems to exacerbate the primacy problem. To the extent that such functions also determine the welfare levels of other cities, redistribution of them among all cities may seriously be considered.

Ⅵ. BIBLIOGRAPHY

1. Alonso, William, "Urban and Regional Imbalances in Economic development", Economic Development and Cultural Change, Vol. 17, No. 1(October 1968), pp. 1-14.
2. Alonso, William, "The Economics of City Size", Papers of the Regional Science Association, Vol. 26, 1971, pp. 67-83.
3. Beckmann, Martin J., "City Hierarchies and the Distribution of City Size", Economic Development and Cultural Change, Vol. 6, 1958, pp. 243-248.
4. Beckmann, Martin J., and John C. McPherson, "City Size Distribution in a Central Place Hierarchy : An Alternative Approach", Journal of Regional Science, Vol. 10, 1970, pp. 25-30.
5. Berry, Brian J.L., "City Size Distributions and Economic development", Economic Development and Cultural Change, Vol. 9, (July 1961), pp. 573-588.
6. Borts, George H., and Jerome L. Stein, Economic Growth in a Free Market, New York : Columbia University Press, 1964.
7. Chipman, John S., "External Economies of Scale and Competitive Equilibrium", Quaterly Journal of Economics, Vol. LXXXIV, No. 3, (August 1970), pp. 347-385.
8. El-Shakhs, Salah, "Development, Primacy, and Systems of Cities", The Journal of Development Areas, Vol. 7, (October 1972), pp. 11-36.
9. Engle, Robert F., "Issues in the Specification of an Econometric Model of Metropolitan Growth", Journal of Urban Economics, 1974, pp. 250-267.
10. Friedmann, John and Clyde Weaver, Territory and Function : The Evolution of Regional Planning, Berkeley and Los Angeles : University of California Press, 1979.
11. Henderson, J. Vernon, "The Sizes and Types of Cities", American Economic Review, Vol. LXIV, No. 4,(September 1974), pp. 640-656.
12. Henderson, J. Vernon, "A Framework for International Comparisons of Systems of Cities", Urban and Regional Report, No. 80-3, The World Bank : (March 1980).
13. Hoover, Edgar M., "Location Theory and the Shoe and Leather Industries", Harvard Economic studies, Vol. LV, Cambridge, MA : Harvard University Press, 1937.
14. Isard, Walter, Location and space economy, Cambridge, MA : MIT Press, 1956.
15. Kemp, Murray C., "The Efficiency of Competition as an Allocator of Resources : I. External Economies of Produc tion", Canadian Journal of Economics and Political Science, Vol. XXI, No. 1, (February 1955), pp. 30-42.
16. Kemp, Murray, C., The Pure Theory of International Trade, Englewood Cliffs, NJ : Prentice-Hall, 1964.
17. Kwon, Young G., "Urbanization, Primacy, and Economic development : An Indonesian Perspective", Cambridge, MA : Department of Urban Studies and Planning, MIT, 1981, mimeograph.

18. Losch, August, The Economics of Location, New Haven, CT : Yale University Press, 1954.
19. Melvin, James R., "Increasing Returns to Scale as a Determinant of Trade", Canadian Journal of Economics, Vol. 2, No. 3,(August 1969), pp. 389-402.
20. Miron, John R., "Migration and Economic Growth", Regional Science and Urban Economics, Vol. 9, 1979, pp. 159-183.
21. Moomaw, Ronald L., Productivity and City Size : A Critique of the Evidence", Quarterly Journal of Economics, Vol. 96, (November 1981), pp. 675-688.
22. Ohlin, Bertil, Interregional and International Trade, Cambridge, MA : Harvard University Pressd, 1933.
23. Richardson, Harry W., "Optimality in City Size, Systems of Cities, and Urban Policy : A Sceptic's View", Urban Studies, Vol.59, No. 1, (February 1971), pp. 29-48.
24. Richardson, Harry W., "National Urban Development Strategies in Developing Countries", Urban Studies, Vol. 18,1981, pp. 267-283.
25. Rodwin, Lloyd, "Metropolitan Policy for Developing Areas", in Lloyd Rodwin ed., The Future Metropolis, New York : George Braziller, 1961, pp. 171-189.
26. Rodwin, Lloyd, Nations and Cities : A Comparison of Strategies for Urban Growth, Boston : Houghton Mifflin, 1970.
27. Rosen, Kenneth and Mitchell Resnick, "The Size Distribution of cities : An Examination of the Pareto Law and Primacy", Journal of Urban Economics, Vol. 8, 1980, pp. 165-186.
28. Schaeffer, Gordon P., "The Urban Hierarchy and Urban Area Production Function : A Synthesis", Urban Studies, Vol. 14, 1977, pp. 315-326.
29. Segal, David, "Are There Return to Scale in City size ? ", Review of Economics and Statistics, Vol. LVIII, 1976, pp. 339-350.
30. Shefer, Daniel, "Localization Economies in SMSA's : A Production Function Analysis", Journal of Regional Science, Vol. 13, No. 1, 1973, pp. 55-64.
31. Sveikauskas, Leo, "The Productivity of Cities", Quarterly Journal of Economics, Vol. LXXXIX, (August 1975), pp. 393-413.
32. Thompson, wilbur R., "Internal and External Factors in the Development of Urban Economies", in Harvey S. Perloff and Lowdon Wingo, Jr., eds., Issues in Urban Economics, Baltimore : Johns Hopkins Press, 1968, pp. 43-62.
33. Thompson, Wilbur R, "The National System of Cities as an Object of Public Policy", Urban Studies, Vol. 9, No. 1, (February 1972), pp. 99-116.
34. United Nations, "1978 World Population Trends and Policies", 1977 Monitoring Report, New York, 1978.
35. United Nations, "1980 World Population Trends and Policies", 1979 Monitoring Report, New York, 1980.
36. Wheaton, William C., and Hisanobu Shishido, "Urban Concentration, Agglomeration Economies, and the Level of Economic De- velopment", Economic Development and Cultural Change, Vol. 30, 1981, pp. 17-30.