CHAPTER VII

OPTIMAL GROWTH OF CITIES

There have been very few works on the mathematical theory of urban growth. Recently, however, this area has begun to attract the attention of theorists. Miyao (1977a,b) analyzed capital accumulation in urban transportation. Rabenau (1976) considered the optimal growth of a small and open factory town with durable housing stock. Fujita (1976a,b) studied accumulation of more than one kind of durable housing capital. These works, however, are concerned only with growth of a certain city, despite the fact that in a modern economy the migration of households and firms is not difficult. Limiting the analysis to a single city prevents us from examining the interaction among cities. Isard and Kanemoto (1976) made an attempt to consider the optimal growth of an economy consisting of many cities and their hinterlands, though the model there is too complicated to go beyond the derivation and interpretation of first order conditions. This motivates the drastic simplifying assumptions of the model in this chapter.

For the first time, productive capital appears in our economy. Like capital in simple neoclassical growth models, it has a number of convenient features: it does not depreciate; it can be applied to any task; and if it is not needed for production, it can be eaten. In addition, because we are considering an economy with many cities, we also require capital that can be moved between and within cities costlessly.

The time dimension must be added to analyze capital accumulation. Since we already have the spatial dimension, the model becomes quite complicated. To keep the model manageable, we make the following drastic simplifying assumptions.

a. The economy consists of cities only: there is no rural sector (except possibly for the constant rural rent).

b. Capital accumulation occurs only in the urban production sector and there is no capital accumulation in the transportation sector.

c. Capital is perfectly mobile: capital can be moved between an within cities instantaneously and without cost.

d. Households are perfectly mobile.

e. All cities are identical. This assumption can be made only when capital is perfectly mobile: otherwise, a new city has zero capital stock initially, and cannot be identical with older cities. Under the assumption of mobility, capital stock in other cities can be instantaneously moved to the new city, and all cities can be made identical.

We assume that there is a Marshallian externality of the sort discussed in Chapter II. At the optimal city size declining production costs, which result from increasing city size, exactly match increasing transportation costs.

The utility level that households achieve in our economy has been determined by the amounts of land and of the consumer good they received. Now that the economy contains capital, some part of output can be invested in physical capital.

The problem of determining the optimal path of our economy may be solved in two stages. At each instant of time, all cities must maintain the optimal spatial allocation, as in Chapter I. The key difference is that when there is capital, the optimization is performed using the part of the product allocated to current consumption, rather than the entire product. The maximum utility level achievable in each city is then obtained as a function, $U(c, P)$, of current consumption, c, and the population of the city, P. In section 1, the model from Chapter I is reformulated with capital, and in section 2 the static spatial optimum is derived given the level of consumption.

At this point the inhabitants of each city know how to allocate their consumption, but not how much of their total product to consume. We assume that they choose to maximize the undiscounted sum of utilities over an infinite time horizon. That is, they are exactly as concerned about the welfare of their most remote descendants as they are about their own. We chose this assumption mainly for the sake of simplicity, but also because we see no moral justification for discounting the welfare of future generations. At any rate, it is quite easy to extend our analysis to the discounted case.

In maximizing the undiscounted sum over an infinite time horizon, we encounter a well-known difficulty: the undiscounted sum of future utilities is infinite, and we are left attempting to compare infinities. Economists have, of course, found several ways of avoiding this problem. In this chapter, we adopt a version of the *Ramsey device* used by Koopmans (1965). This approach changes the origin of the instantaneous utility function, taking the utility level of the *optimal steady state* as zero, where the optimal steady state, or the Golden Rule, is the balanced growth path which maximizes the utility level among all feasible balanced growth paths. If $u(x)$ denotes the utility level at time t, and u^* that at the optimal steady state, the sum of the difference over infinite time horizon,

$$\int_0^\infty [u(t) - u^*]dt,$$

is maximized. The new objective function turns out to be bounded from above, and the difficulty of comparing infinities disappears. In section 3, the objective function is maximized with respect to the paths of consumption and population of a city.

In optimal growth theory, it is usually assumed that the utility function is concave. In our model, however, the maximum utility function, $U(c, P)$, may not be concave, although the original utility function over the consumer good and land is assumed to be concave. As it turns out, the maximum utility function is not even quasi-concave in most cases. This does not create serious difficulties for our analysis, if we assume that the concavity of the per capita production function is strong enough.

In section 4, a phase diagram analysis is carried out to determine whether a city grows during the process of capital accumulation. Section 5 contains remarks on the limitations of the model, and speculations on how the results might be modified if the model is extended.

1. The Model

Consider the growth of an economy consisting of cities. Let capital accumulation occur in the urban production sector and the number of cities change in the process of growth. Assume there is no non-urban sector and the total population of the whole economy is partitioned into cities. This assumption is clearly unrealistic and precludes the analysis of the evolution of an economy through different stages, for example, from the rural stage to the urban stage, as analyzed by Isard and Kanemoto (1976). Considering the complexity of the problem and the dominance of the urban sector in a modern economy, however it seems worthwhile to start with this simple formulation.

As discussed in Chapter II, economic factors which cause cities can be classified into three categories: concentration of immobile factors, increasing returns to scale, and externalities. In this chapter we consider cities based on a Marshallian externality of the kind analyzed in section 5 of Chapter II. Instead of starting from the production function of an individual firm, we simply assume that the aggregate production function of a city can be written as

$$F(P,K,P), \tag{1.1}$$

where P and K are respectively the population and the aggregate capital stock of the city. The production function is homogeneous of degree one with respect to the first two terms and the derivative with respect to the third term is positive. The production function, therefore, exhibits increasing returns to scale if the third term is taken into account.[1]

Because it is easier to work with the capital-labour ratio and consumption per capita, $k = K/P$ and c, than with the absolute quantities, we want a per capita production function $f(k,P)$. By the homogeneity assumption, the per capita production function is

$$f(k,P) = F(1,k,P) = F(1,K/P,P), \tag{1.2}$$

where

$$f_P > 0. \tag{1.3}$$

We assume that the per capita production function is strictly concave and

$$f_k > 0. \tag{1.4}$$

As in previous chapters, we assume that all cities are identical. If at time t the population of the whole economy is $\overline{P}(t)$, the capital stock for the whole economy is $\overline{K}(t) = \overline{P}(t)k(t)$, and per capita consumption of the produced good is $c(t)$, then the output available for capital accumulation after consumption is

$$\overline{k}(t) = \overline{P}(t)\,f(k(t),P(t)) - \overline{P}(t)c(t). \tag{1.5}$$

---

[1] It is not difficult to show that if an individual firm has a production function $\tilde{f}(\ell,k,P)$ where ℓ and k are respectively labour and capital inputs, the aggregate production function can be written as (1.1) when the number of firms is optimal.

If we assume that the population growth rate is a constant

$$l = \overline{P}(t)/\overline{P}(t), \tag{1.6}$$

then (1.5) can be rewritten as

$$\dot{k}(t) = f(k(t), P(t)) - l\,k(t) - c(t). \tag{1.7}$$

The spatial structure of a city is the same as in the previous chapters. $q(x)dx$ of land is available in the ring between x and $x + dx$, where x is the distance from the center of a city. A household living at x at time t has a lot size $h(x,t)$. Then there are $(q(x)/h(x,t))dx$ households between x and $x + dx$ at time t. A household at x at time t consumes $z(x,t)$ of the produced good and spends $T(x)$ on commuting costs expressed in terms of the produced good. For simplicity, we assume that there is no capital accumulation or no technological progress in the transportation sector. Note that we have changed the notation for commuting costs and that t now denotes time. A city uses $c(t)P(t)$ of the produced good for consumption, which includes direct consumption, commuting costs, and the payment of the rural rent R_a. The *resource constraint* for a city is then

$$c(t)P(t) = \int_0^{\bar{x}(t)} \{[z(x,t) + T(x)]/h(x,t) + R_a\} q(x)dx, \tag{1.8}$$

where $\bar{x}(t)$ is the edge of the city at time t.

The *population constraints* are

$$P(t) = \int_0^{\bar{x}(t)} [q(x)/h(x,t)]dx, \tag{1.9}$$

and

$$\overline{P}(t) = n(t)P(t), \tag{1.10}$$

where $n(t)$ is the number of cities at time t. We shall ignore the constraint that $n(t)$ be an integer and take $n(t)$ as a continuous variable.

The utility function is $u(z,h)$, and we impose the constraint that the utility level be equal everywhere at each instant of time. The utility level may vary over time.

The *equal-utility constraint* can be written

$$u(t) = u[z(x,t), h(x,t)]. \tag{1.11}$$

Having set up the model, our problem is to maximize the undiscounted sum over an infinite time horizon:

$$\int_0^\infty [u(t) - u^*]dt, \tag{1.12}$$

subject to the constraints (1.7) through (1.11), and the initial condition,

$$k(0) = k_0, \tag{1.13}$$

where u^* is the utility level in the optimal steady state. The problem is solved in two stages.

2. Optimal Spatial Structure

In this section the first stage optimization is carried out for given $c(t)$ and $P(t)$, and the properties of the maximum utility function $U(c(t), P(t))$ are examined. This problem is exactly the same as the one in subsection 2.1 of Chapter I if we substitute $c(t)$, $P(t)$, and $T(x)$ for w, P, and $t(x)$. The utility level is maximized under the resource constraint, the population constraint, and the equal-utility constraint, which are in this case, (1.8), (1.9), and (1.11) respectively. Control variables are $z(x)$ and $h(x)$, and control parameters are $\bar{x}$ and u. The time variable t is suppressed in this section, since it plays no role in the optimization.

The first order conditions can be rewritten

$$u_h(z(x), h(x))/u_z(z(x), h(x)) = R(x), \tag{2.la}$$

$$y = z(x) + R(x)h(x) + T(x), \tag{2.1b}$$

$$R(\bar{x}) = R_a, \tag{2.1c}$$

after simple manipulations. As in Chapter I, the optimal solution can be achieved as a competitive equilibrium if all households receive the same income y. The solution, therefore, can be described by using the bid rent function $R(I(x), u) = R(y - T(x), u)$ defined in Equation (1.12) of Chapter I:

$$y = c + s, \tag{2.2a}$$

$$sP = \int_0^{\bar{x}} [R(y - T(x), u) - R_a] q(x) dx, \tag{2.2b}$$

$$P = \int_0^{\bar{x}} R_I (y - T(x), u) q(x) dx, \tag{2.2c}$$

$$R(y - t(\bar{x}), u) = R_a. \tag{2.2d}$$

s is the social dividend each household receives and is equal to the total differential rent divided by the population of the city. (2.2a) and (2.2b) correspond to (1.28) in Chapter I. (2.2c) is a restatement of the population constraint using the property of the bid rent function: $R_I = 1/h$.

If c and P are given, the four equations, (2.2a)-(2.2d), determine the four variables, y, s, $\bar{x}$, and u. The utility level which is obtained can then be written as a function $U(c, P)$ Total differentiation of (2.2) yields the partial derivatives of the maximum utility function:

$$U_P(c, P) = s / \int_0^{\bar{x}} R_u q(x) dx < 0, \tag{2.3}$$

$$U_c(c, P) = -P / \int_0^{\bar{x}} R_u q(x) dx > 0, \tag{2.4}$$

where subscripts P and c denote partial derivatives with respect to P and c respectively. Thus an increase in the population of a city, given the consumption of resources per capita, lowers the utility level which can be attained in the city. An increase in per capita consumption given the population raises the utility level. The marginal rate of substitution between P and c is equal to the negative of the social dividend divided by the population:

$$S(c, P) \equiv U_P(c, P) / U_c(c, P) = -s / P. \tag{2.5}$$

Further properties of the maximum utility function are difficult to derive in the general case. The following results for four cases have been obtained by tedious calculations. The cases are

(i) the Leontief utility function $u(z, h) = [\min(z/a, h)]^{1/g}, g > 1$, in a linear city, $q(x) = q$;

(ii) the Leontief utility function in a pie-slice city, $q(x) = qx$;

(iii) the Cobb-Douglas utility function $u(z,h)=(z^{a}h^{1-a})^{1/g}$, $g>1$, in a linear city; and

(iv) the Cobb-Douglas utility function in a pie-slice city.

In all cases, linear commuting costs are assumed: $T(x)=Tx$. The results are

(1) U_{cc} is negative in the linear city cases (i) and (iii).
In circular cities (ii) and (iv), U_{cc} is positive if g is close to 1 and is negative for a large enough g (in case (iv) we have proven this only in the case $R_a=0$).

(2) U_{PP} is positive in all cases (in the Cobb-Douglas cases we have proven this only in the case of $R_a=0$). This shows that $U(c,P)$ is not concave.

(3) $U(c,P)$ is not usually quasi-concave. In order for U to be quasi-concave, $\Delta=2U_{cP}U_cU_P-U_{cc}U_P^{2}-U_{PP}U_c^{2}$ must be nonnegative. In the case of the Leontief utility function, Δ equals zero in a linear city, and Δ is negative if $c>T\bar{x}$ in a circular city. In the Cobb-Douglas case, Δ is negative in a linear city and in the case of $R_a=0$ in a circular city.

(4) $S_c(c,P)$ is negative in all cases. This implies that -*P* would be a normal good if $U(c,P)$ were quasi-concave. (Note that P is a 'bad' and hence $-P$ is a good.)

(5) $S_P(c,P)$ is positive in all cases (in case (iv) we have proven this only in the case of $R_a=0$). This implies that c would be an inferior good if $U(c,P)$ were quasi-concave.

These results show that even if the original utility function $u(z,h)$ is concave, the maximum utility function $U(c,P)$ is not usually well behaved: $U(c,P)$ is usually neither concave nor quasi-concave. As it turns out, however, this does not cause a serious difficulty in the second stage optimization if the concavity of the production function (1.2) is strong enough.

3. Optimal Growth of Cities

In the second stage of our optimization procedure, the undiscounted sum over an infinite time horizon,

$$\int_0^\infty [U(c(t), P(t)) - u^*]dt, \tag{3.1}$$

is maximized subject to (1.7), (1.10), and (1.13). $U(c, P)$ is the maximized utility level from section 2 and u^* is the utility level at the optimal steady state. Since $n(t)$ appears only in the constraint (1.10) and is taken as a continuous variable, the problem is equivalent to the one of maximizing (3.1) under the constraints (1.7) and (1.13) with respect to $c(t)$ and $P(t)$. Although the population of a city $P(t)$ must be greater than one, we ignore this constraint, assuming that it is always satisfied along the optimal path.

Before solving this problem, we first examine the optimal steady state, at which the utility level is maximized among all feasible steady states. The optimal steady state is therefore the solution to the problem of maximizing

$$U(c, P)$$

subject to

$$f(k, P) - l\,k - c = 0, \tag{3.2}$$

with respect to c, P, and k.

First order conditions for an interior optimum are

$$f_k(k, P) = l\,, \tag{3.3a}$$

$$\frac{U_P(c, P)}{U_c(c, P)} + f_P(k, P) = 0\,. \tag{3.3b}$$

The first equation is the usual condition that the system operate at the biological rate of interest: the marginal productivity of capital must equal the population growth rate. The second equation requires that the population of a city be determined so that the *per capita* marginal external benefit on the production side equals the marginal rate of substitution between population and resource consumption per capita. From (2.5) and (2.2b), this is equivalent to

$$P[Pf_P(k, P)] = \int_0^{\bar{x}} [R(x) - R_a]q(x)dx, \tag{3.4}$$

which may be interpreted as the condition obtained in Chapter II that the total differential rent equals the total Pigouvian subsidy. An additional worker in a city produces $f(k, P)$ of the product himself, but at the same time increases the population of the city and raises the production of other workers by Pf_P. The latter is the marginal external benefit, and the Pigouvian subsidy must equal Pf_P to achieve an efficient allocation. The left side of (3.4) is, therefore, the total amount of the Pigouvian subsidy in the city, which must equal the total differential rent when the number of cities is optimal.

The second order conditions are as follows.

$$f_{kk} \le 0, \tag{3.5a}$$

$$S_P - S \bullet S_c + f_{PP} - (f_{kP})^2 / f_{kk} \le 0, \tag{3.5b}$$

where $S(c,P)$ is the marginal rate of substitution between P and c and is defined in (2.5). The first two terms are

$$\begin{aligned} S_P - S \bullet S_c &= -\frac{1}{U_c^3}\left[2U_{cP}U_cU_P - U_{cc}U_P^2 - U_{PP}U_c^2\right] \\ &\equiv -\Delta / U_c^3. \end{aligned}$$

Since $U(c,P)$ is not usually quasi-concave as seen in section 2, $S_P - S \bullet S_c$ is usually positive. $f_{PP} - (f_{kP})^2 / f_{kk}$ is, however, negative if $f(k,P)$ is concave. (3.5b) can, therefore, be satisfied if the concavity of the production function is strong enough. (3.5a) is satisfied because we assumed that the production function is concave. We henceforth assume that (3.5a) and (3.5b) are satisfied with strict inequalities. We also assume that the optimal steady state is unique.

The following two observations can be immediately obtained from the first order conditions (3.3). First, unlike usual one sector growth models, the optimal steady state depends on the shape of the utility function. The population of a city serves as a link between the consumption side and the production side, and the capital-labour ratio at the optimal steady state is affected by the shape of the utility function. Second, at the optimal steady state, the configuration of a city remains exactly the same, and the number of cities increases at the same rate as the population growth.

Now, let us go back to the original problem of maximizing (3.1) with respect to $c(t)$ and $P(t)$ subject to (1.7) and (1.14). As shown in section 2 of the appendix on optimal control, the Hamiltonian for this problem is

$$\Phi = U(c(t), P(t)) + q(t)[f(k(t), P(t)) - l\,k(t) - c(t)], \tag{3.6}$$

where $q(t)$ is an adjoint variable associated with the constraint (1.7). $q(t)$ satisfies the adjoint equation:

$$-\dot{q}(t) = q(t)[f_k(k(t), P(t)) - l], \tag{3.7}$$

and the Hamiltonian must be maximized with respect to $c(t)$ and $P(t)$. The first order conditions for the maximization are

$$U_c(c(t), P(t)) = q(t), \tag{3.8a}$$

$$U_P(c(t), P(t)) + q(t) f_P(k(t), P(t)) = 0, \tag{3.8b}$$

and the second order conditions are

$$U_{cc} \le 0, \tag{3.9a}$$

$$U_{PP} + qf_{PP} \le 0, \tag{3.9b}$$

$$U_{cc}U_{PP} - (U_{cP})^2 + qU_{cc} f_{PP} \ge 0. \tag{3.9c}$$

As seen in section 2, (3.9a) is satisfied if the concavity of the original utility function $u(z,h)$ is strong enough. For (3.9b) to be satisfied, f_{PP} must be negative and its absolute value must be greater than U_{PP}/q, since U_{PP} is usually positive, and by (3.8a), q is also positive. In (3.9c) the sum of the first two terms is usually negative. Again, f_{PP} must be negative with a large absolute value.

Combining (3.7) and (3.8a) yields the differential equation:

$$U_{cc}\dot{c}(t) + U_{cP}\dot{P}(t) = U_c[l - f_k], \tag{3.10}$$

and from (3.8a) and (3.8b) we obtain

$$S(c(t), P(t)) + f_P(k(t), P(t)) = 0. \tag{3.11}$$

Using (2.5), (3.11) becomes

$$s(t) = P(t) f_P(k(t), P(t)). \tag{3.12}$$

Thus the social dividend equals the Pigouvian subsidy at each point of time along the optimal path. In other words, the total amount of the Pigouvian subsidy for residents of the city must always equal the total differential rent.

Since there is no constraint on $k(t)$ at terminal time $t = \infty$, the transversality condition must be obtained to determine the value of $k(t)$ at $t = \infty$. If we can show that the optimal path converges to the optimal steady state, the transversality condition must be

$$\lim_{t \to \infty} q(t)k(t) = q^* k^*, \tag{3.13}$$

where $q^* = U_c(c^*, P^*)$, and asterisks denote the optimal steady state values of the variables.

We prove that the optimal path converges to the optimal steady state in two steps.[2] In the rest of this section, we show that the optimal path visits any arbitrarily small neighbourhood of the optimal steady state. This result still allows the possibility that the optimal path enters a neighbourhood of the optimal steady state but leaves there eventually. In the next section, we examine the behaviour of the optimal path near the steady state, and show that the steady state is a saddle point. Since this means that all paths except the one which converges to the saddle point diverge, the only path that visits an arbitrarily small neighbourhood of the optimal steady state is the convergent one. Thus the optimal path must converge to the optimal steady state, and (3.13) is in fact the required transversality condition.

To establish that the optimal path must visit an arbitrary neighbourhood of the steady state, we observe that the Kuhn-Tucker Theorem shows that when the constraint qualification is satisfied[3], there exists a multiplier q^* such that the optimal steady state maximizes the Lagrangian

$$U(c, P) + q^*\left[f(k, P) - lk - c\right].$$

Thus the optimal steady state (k^*, c^*, P^*) satisfies

---

[2] This approach is similar to the one used by Scheinkman (1976) in the discounted case with many stocks.

[3] See, for example, Mangasarian (1969). See also section 3 in the appendix on optimal control theory for the explanation of constraint qualification.

$$U(c^*,P^*) + q^*\left[f(k^*,P^*) - lk^* - c^*\right] \\ \geq U(c,P) + q^*\left[f(k,P) - lk - c\right], \quad \text{for any } k, c, \text{and } P. \tag{3.14}$$

It is not difficult to show that under some regularity conditions this inequality can be strengthened to the following: if $|k - k^*| > e$ for any positive e, then there exists $r > 0$ such that

$$u^* = U(c^*,P^*) + q^*\left[f(k^*,P^*) - lk^* - c^*\right] \\ > U(c,P) + q^*\left[f(k,P) - lk - c\right] + r, \quad \text{for any } c \text{ and } P. \tag{3.15}$$

If a path does not visit an arbitrarily small neighbourhood of the optimal steady state, there exists some $e > 0$ such that $|k(t) - k^*| > e$ for any t. Inequality (3.15) then holds for any t and we can integrate it from 0 to ∞ to obtain

$$\int_0^\infty \left[U(c,P) - u^*\right]dt < -q^*\left[k(\infty) - k_0\right] - \int_0^\infty r\,dt. \tag{3.16}$$

Since $k(\infty) \geq 0$, the right side of the inequality is minus infinity. Thus the value of the criterion function of any path that does not visit an arbitrarily small neighbourhood of the optimal steady state is minus infinity.

Now it is easy to construct feasible paths which have values of the criterion greater than $-\infty$. For example, consider a path which approaches the optimal steady state with a constant $\dot{k}(\neq 0)$ and stops there. Such a path always exists if the initial capital-labour ratio, k_0, is larger than k^*, since we can determine $c(t)$ in such a way that $\dot{k}(t)$ is negative and constant until $k(t)$ reaches k^*. Even if the initial capital-labour ratio is smaller than k^*, such a path exists as long as $f(k - P) - lk$ is positive for any k between k_0 and k^*.

Since $\dot{k}$ is constant and is not equal to zero, k^* will be reached within a finite length of time. The value of the criterion up to that time is then finite, and after that time the value can be made equal to zero by setting $c(t) = c^*$ and $P(t) = P^*$. Thus there exists a feasible path with a finite value of the criterion, and any path that does not visit an arbitrarily small neighbourhood of the optimal steady state cannot be optimal.

4. Phase Diagram Analysis

Now we examine the local behaviour of the optimal path near the optimal steady state. The optimal path satisfies differential equations (1.7) and (3.10), and equation (3.11) which must hold at each instant of time. The dynamic system contains three variables: k, c, and P. In order to work with a two-dimensional phase diagram, we use (3.11) to express c as a function, $c(k,P)$, of k and P, and obtain differential equations of k and P. Then implicit differentiation of (3.11) yields derivatives of $c(k,P)$:

$$c_k(k,P) = -f_{Pk}/S_c, \tag{4.1}$$

$$c_P(k,P) = -(S_{PP} + f_{PP})/S_c. \tag{4.2}$$

Observing

$$\dot{c}(t) = c_k \dot{k} + c_P \dot{P},$$

we can rewrite (3.10) as follows using (4.1), (4.2) and (1.7):

$$\dot{P}(t) = \frac{1}{D(k,P)}\{f(k,P)[l - f_k(k,P)] + j(k,P)[f(k,P) - lk - c(k,P)]\}, \tag{4.3}$$

where

$$D(k,P) = U_{cc}(S_P + f_{PP}) - U_{cP}S_c, \tag{4.4a}$$

$$f(k,P) = -U_c S_c, \tag{4.4b}$$

$$j(k,P) = -U_{cc} f_{Pk}. \tag{4.4c}$$

The differential equation (1.7) can also be rewritten using $c(t) = c(k(t), P(t))$,

$$\dot{k}(t) = f(k,P) - lk - c(k,P). \tag{4.5}$$

(4.3) and (4.5) describe the paths that $k(t)$ and $P(t)$ must follow. The optimal steady state is the rest point of (4.3) and (4.5) since (3.2), (3.3a), and (3.3b) hold at the rest point.

To construct the phase diagram, we must know the signs of D, f and j. By simple manipulations, D becomes as follows.

$$D(k,P) = U_{cc} f_{PP} + \left[U_{cc}U_{PP} - (U_{cP})^2\right]/U_c$$
$$\geq 0,$$

(4.6)

where the inequality is obtained from (3.8a) and (3.9c). In order to determine the signs of f and j, we assume

$$f_{Pk} \geq 0, \tag{4.7}$$

$$S_c \leq 0. \tag{4.8}$$

The first inequality implies that capital and population are complementary in production. As mentioned in section 2, the second inequality is satisfied in all the examples we have calculated. Since $U_{cc} \leq 0$ from the second order condition (3.9a), f and j are both nonnegative under these assumptions:

$$f(k,P) \geq 0, \tag{4.9}$$

$$j(k,P) \geq 0. \tag{4.10}$$

These assumptions also imply that

$$c_k(k,P) > 0. \tag{4.11}$$

We now construct the phase diagram of (4.3) and (4.5). Following the usual procedure, we first examine the loci of $\dot{P}=0$ and $\dot{k}=0$. The locus of $\dot{k}=0$ is

$$f(k,P) - lk - c(k,P) = 0, \tag{4.12}$$

and the slope of the locus is

$$\left.\frac{dk}{dP}\right|_{\dot{k}=0} = \frac{f_P - c_P}{c_k + l - f_k}. \tag{4.13}$$

Since by (3.3a) we have $f_k = l$ at the optimal steady state, the slope there is

$$\left.\frac{dk}{dP}\right|_{\dot{k}=0} = \frac{f_P - c_P}{c_k}. \tag{4.14}$$

The denominator is positive by (4.11). By (4.2) the numerator is

$$f_P - c_P = \frac{1}{S_c}\left[f_{PP} + S_P - S \cdot S_c\right],$$

which is also positive from (4.6), (3.5b) and the strict concavity of $f(k,P)$. Thus the $\dot{k}=0$ locus is upward sloping at the Golden Rule:

$$\left.\frac{dk}{dP}\right|_{\dot{k}=0} > 0 \qquad \text{at} \qquad f_k = l. \tag{4.15}$$

Since we have

$$\frac{\partial}{\partial P}\left[f(k,P)-lk-c(k,P)\right]=f_P-c_P>0, \tag{4.16}$$

$\dot{k}=f-lk-c$ is negative above the $\dot{k}=0$ locus and positive below the locus as illustrated in Figure 1.

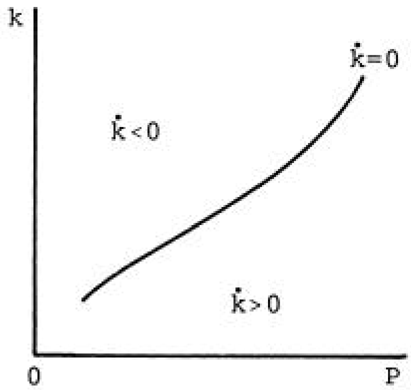

Figure 1
The Locus of $\dot{k}=0$

Next, consider the locus of $\dot{P}=0$. From (4.3) it is a combination of the $\dot{k}=0$ locus and the locus of

$$l-f_k(k,P)=0. \tag{4.17}$$

The slope of the locus of (4.17) is

$$\left.\frac{dk}{dP}\right|_{f_k=l}=-f_{kP}/f_{kk}>0, \tag{4.18}$$

where the inequality is the result of concavity and complementarity. The locus of (4.17) is, therefore, upward sloping. Since

$$\frac{\partial}{\partial k}\left[l-f_k(k,P)\right]=-f_{kk}$$
$$>0,$$

$l-f_k$ is positive above the locus of $l=f_k$ and negative below the locus. This is illustrated in Figure 2.

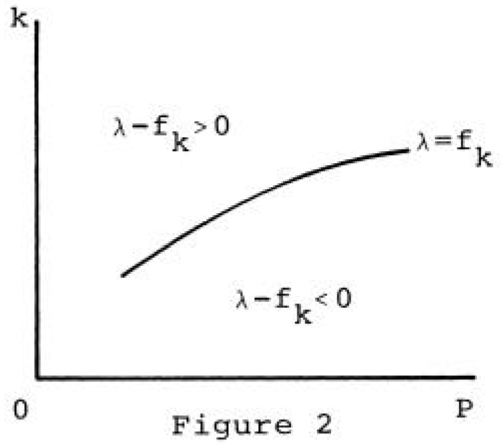

Figure 2
The Locus of $\lambda = f_k$

The locus of $l = f_k$ intersects with the $\dot{k} = 0$ locus at the Golden Rule. The $\dot{k} = 0$ locus is steeper than the locus of $l = f_k$ at the intersection point since the following inequality holds there:

$$\begin{aligned} &\frac{dk}{dP}\Big|_{\dot{k}=0} - \frac{dk}{dP}\Big|_{l=f_k} \\ &= \frac{f_P - c_P}{c_k} + \frac{f_{kP}}{f_{kk}} \\ &= -\frac{S_P - SS_c + f_{PP} - (f_{kP})^2 / f_{kk}}{f_{Pk}} \\ &\geq 0, \end{aligned} \tag{4.20}$$

where we used (3.5b) and (4.7). Figure 3 illustrates the relationship between the two loci. Since *D*, *f* and *j* are all nonnegative, the $\dot{P} = 0$ locus passes through regions (A) and (C) in Figure 3, and $\dot{P}$ is positive on the side of region (D). There are three possibilities:

(i) the $\dot{P} = 0$ locus is downward sloping,

(ii) the $\dot{P} = 0$ locus is upward sloping but flatter than the $l = f_k$ locus, and

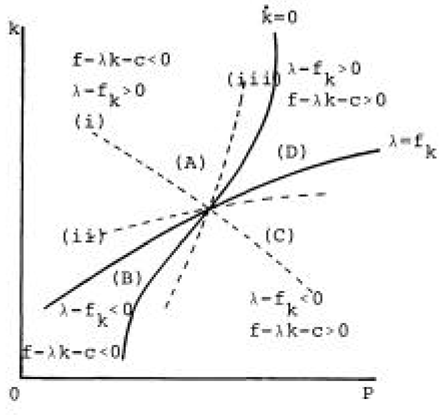

Figure 3. The Loci of $\dot{k}=0$, $\lambda=f_k$, and $\dot{P}=0$

(iii) the $\dot{P}=0$ locus is upward sloping and steeper than the $\dot{k}=0$ locus.

The slope of the $\dot{P}=0$ locus is, at the Golden Rule,

$$\left.\frac{dk}{dP}\right|_{\dot{P}=0}=\frac{Df_{kP}}{U_{cc}f_{kk}[(-1/f_{kk})(f_{pk})^2-(-U_c/U_{cc})(S_c)^2]}$$

$$\gtrless 0 \qquad \text{as} \qquad (-1/f_{kk})(f_{Pk})^2 \gtrless (-U_c/U_{cc})(S_c)^2$$

(4.21)

and

$$\left.\frac{dk}{dP}\right|_{\dot{P}=0}-\left.\frac{dk}{dP}\right|_{\dot{k}=0}$$

$$=\frac{U_c(S_c)^2[S_P-SS_c+f_{PP}-(f_{kP})^2/f_{kk}]}{U_{cc}f_{Pk}[(-1/f_{kk})(f_{kP})^2-(-U_c/U_{cc})(S_c)^2]}$$

$$\gtrless 0 \qquad \text{as} \qquad (-1/f_{kk})(f_{Pk})^2 \gtrless (-U_c/U_{cc})(S_c)^2.$$

(4.22)

These relationships imply that

$$\left.\frac{dk}{dP}\right|_{\dot{P}=0}>\left.\frac{dk}{dP}\right|_{\dot{k}=0} \qquad \text{if} \qquad (-1/f_{kk})(f_{Pk})^2>(-U_c/U_{cc})(S_c)^2 \qquad (4.23)$$

$$\frac{dk}{dP}\bigg|_{\dot{P}=0} < 0 \qquad \text{if} \qquad (-1/f_{kk})(f_{Pk})^2 < (-U_c/U_{cc})(S_c)^2 \tag{4.24}$$

Thus case (i) is obtained if $(-1/f_{kk})(f_{Pk})^2 < (-U_c/U_{cc})(S_c)^2$, and case (iii) otherwise, but case (ii) never occurs.

In case (i), we obtain a phase diagram depicted in Figure 4. The optimal steady state is a saddle point and all paths except for the two stable branches diverge. Since it was shown in the preceding section that the optimal path must visit any arbitrarily small neighbourhood of the optimal steady state, the optimal path must be one of the stable branches. The diagram also shows that at least in the neighbourhood of the steady state the optimal path is either in the region where $\dot{k} > 0$ and $\dot{P} < 0$ or in the region where $\dot{k} < 0$ and $\dot{P} < 0$. *The population of a city therefore declines as capital accumulates*. Notice,

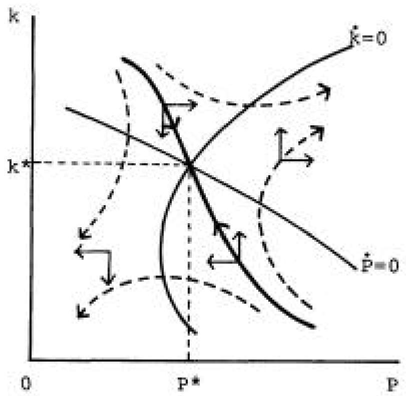

Figure 4. A Phase Diagram

however, that this conclusion may not hold globally as Figure 5 illustrates.

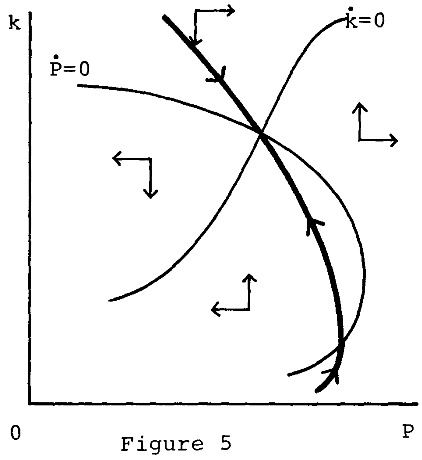

Figure 5
Global Behaviour Which Is Different from Local Behaviour

In case (iii), we obtain a phase diagram like Figure 6. The optimal steady state is a saddle point in this case as well and the optimal solution is either of the stable branches. It can be seen from the diagram that the population of a city rises as capital accumulates in the neighbourhood of the Golden Rule.

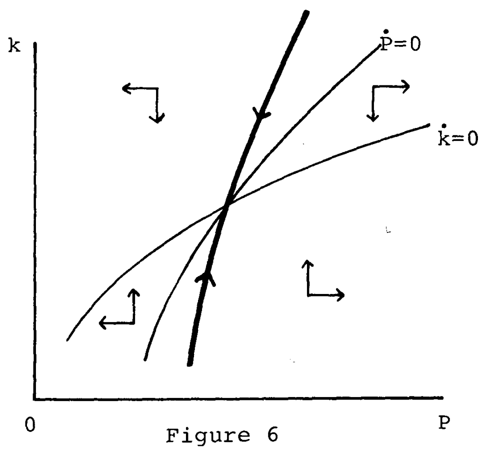

Figure 6
An Upward Sloping Ṗ=0 Locus

These results are summarized in the following theorem:

Theorem 1: Suppose $S_c \le 0$ *and* $f_{Pk} \ge 0$. *If* $(-1/f_{kk})(f_{Pk})^2 < (-U_c/U_{cc})(S_c)^2$, *then the population of a city falls as capital accumulates in the neighbourhood of the optimal* *steady state ; and if* $(-1/f_{kk})(f_{Pk})^2 > (-U_c/U_{cc})(S_c)^2$ *, then the population rises.*

The assumption of complementarity, $f_{Pk} \geq 0$, is rather arbitrary although the assumption is satisfied in most widely used production functions such as the Cobb-Douglas and CES functions. If capital and population are anticomplementary, the following result is obtained.

Theorem 2: Suppose $S_c \leq 0$ *and* $f_{Pk} \leq 0$. *Then the population of a city falls as capital accumulates in the neighbourhood of the optimal steady state.*

Labour augmenting technical progress, or Harrod neutral technical progress, can be incorporated in this analysis quite easily although other types of technical progress are not easy to handle. When the rate of labour augmenting technical progress is s, the same result as in the case without the technical progress is obtained if l is replaced by $l + s$ and P by the population in terms of efficiency labour, $Q = Pe^{st}$. Since P and Q have the relationship:

$$\frac{\dot{P}}{P} = \frac{\dot{Q}}{Q} - s , \qquad (4.25)$$

the rate of increase of the population of a city is smaller by the technical progress rate than the case without the technical progress. Thus labour augmenting technical progress introduces a tendency for city size to decline over time.

The reason why the sign and the magnitude of f_{Pk}, are crucial in Theorems 1 and 2 must be obvious. The population size is determined in such a way that $S + f_P = 0$, i.e., the marginal cost of having a bigger population on the consumption side balances the marginal externality benefit on the production side. If f_{Pk}, is positive, an increase in the capital-labour ratio increases the marginal benefit and tends to increase the population of a city. This tendency would be offset if the marginal cost on the consumption side rises. As capital-labour ratio rises, per capita consumption usually increases. If S_c is nonpositive as assumed in the Theorems, the marginal cost 2 also rises. Theorem 1 states that when $(-U_c/U_{cc})(S_c)^2$ is greater than $(-1/f_{kk})(f_{Pk})^2$, this effect overwhelms the effect of the rise in the marginal benefit. If f_{Pk}, is negative, both effects work in the same direction and the population of a city always declines in the process of capital accumulation as in Theorem 2.

5. Concluding Remarks

We have characterized the condition required for capital accumulation to be accompanied by an increase in the population of a city. It was shown that the population growth tends to occur if capital and the external economy are complementary in production and that the population tends to decline if the marginal rate of substitution between the population and consumption becomes greater in absolute value as the consumption increases. It is believed that ordinary factors of production

are usually complementary, although it is not clear whether this is true if there are externalities. In examples that we have calculated, the marginal rate of substitution between the population of a city and consumption rises in absolute value as the consumption increases. Empirical studies are therefore necessary to determine whether capital accumulation favours bigger cities.

It is quite obvious that our model is too simple to capture the complexity of modern cities. It does not deal with the following important aspects of real cities.

First, we do not have a hierarchy of cities. Rather, our cities are identical. More than one kind of good has to be introduced to obtain a hierarchy of cities.

Second, the production function is assumed to remain the same over time (except for the possibility of labour augmenting technical change). It might have been shifting to increase the benefits of bigger cities.

Third, perfect mobility and malleability of capital is not a realistic assumption, and there are costs involved in building a new city, which tends to reduce the number of cities and hence to increase the size of a city.

Fourth, there is a good reason to believe that a market economy has a very different growth path from the optimal one. As shown in Chapter II, the market equilibrium is not unique and a city size greater than the optimum may well be an equilibrium.

Fifth, technical progress and capital accumulation in transportation sector has worked to reduce the cost of bigger cities.

REFERENCES

Fujita, M., (1976a), "Spatial Patterns of Optimal Growth: Optimum and Market, " *Journal of Urban Economics* 3, 193 -208.

Fujita, M., (1976b), "Toward a Dynamic Theory of Urban Land Use, " *Papers of Regional Science Association* 37, 133-165.

Isard, W. and Y. Kanemoto, (1976), "Stages in Space-Time Development, " *Papers of Regional Science Association* 37, 99-131.

Koopmans, T.C., (1965), "On the Concept of Optimal Economic Growth, " in: *Study Week on Econometric Approach to Planning, Pontificiae Scientiarum Scripta Varia* XXVIII, (Rand McNally, Chicago).

Mangasarian, O.L., (1969), *Nonlinear Programming*, (McGraw-Hill, New York).

Miyao, T., (1977a), "The Golden Rule of Urban Transportation Investment," *Journal of Urban Economics* 4, 448-458.

Miyao, T., (1977b), "A Long-Run Analysis of Urban Growth Over Space, " *Canadian Journal of Economics* 10, 678-686.

Rabenau, B., (1976), "Optimal Growth of a Factory Town, " *Journal of Urban Economics* 3, 97-112.

Ramsey, F.P., (1928), "A Mathematical Model of Saving, " *Economic Journal* 38, 543-559.

Scheinkman, J.A., (1976), "On Optimal Steady States of n-Sector Growth Models when Utility is Discounted, " *Journal of Economic Theory* 12, 11-30.