Optimum Development in a Dual Economy
Author(s): N. H. Stern
Source: *The Review of Economic Studies*, Vol. 39, No. 2 (Apr., 1972), pp. 171-184
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/2296869
Accessed: 09/09/2013 07:00

# Optimum Development in a Dual Economy[1,2]

N. H. STERN
*St Catherine's College, Oxford*

## I. INTRODUCTION

There has been considerable interest in recent years in the development of models to illuminate the problems of optimum investment and employment in the advanced sector of under-developed countries. The labour for the advanced sector is assumed to be provided by the traditional sector. Sen [14] has shown that decisions for the two policies must be taken together and Marglin [11] and Dixit [4] have developed optimum growth models to indicate how these decisions might be made. All three writers have adopted the Lewis [8] assumption of a perfectly elastic supply of labour from the traditional sector at an institutionally fixed wage.

An important objection to the analysis of Marglin and Dixit is the inappropriate specification of the objective function. In optimum growth models that are not designed for dual economies it is customary to use as an objective function $\int_0^\infty u(C/N)e^{-rt}dt$, where $C$ is aggregate consumption, $N$ is total population, $r$ is a discount rate and $u$ is the instantaneous valuation function. With the usual assumption of exponential population growth this formulation includes weighting the instantaneous valuation function by the current population. The implicit assumption, therefore, is that the government can adjust the intra-generational distribution of income to its own liking or that this distribution remains fixed. This assumption, of course, cannot be relevant for the analysis of a dual economy. Indeed, that it is not relevant is part of the essence of the problem since the development of the dual economy is seen as the transfer of the population from one sector to the other, and consumption per head in the advanced sector is usually higher than that in the traditional sector. However Marglin and Dixit apply just such a valuation function to their models where $C$ is the consumption out of the output of the advanced sector. Thus output from the traditional sector is not valued and the important intra-generational distribution consequences of employment policy are ignored.

The valuation function used here is the natural extension of the usual Benthamite valuation $Nu(C/N)$ to intra-generational distribution. We take as the instantaneous valuation function[3] $Lu(C/L)+(N-L)u(\alpha)e^{-rt}$ where $N$ is population, $C$ is consumption out of advanced sector output, $L$ is employment in the advanced sector and $\alpha$ is traditional sector consumption per head. In this formulation, therefore, the special costs of urban living are ignored, but this could be dealt with by deflating $C$ to allow for the higher urban cost of living. We assume $u$ is concave.

Output in the advanced sector is $f(K, L)$ where $K$ is the capital stock in the advanced

[1] *First version received April* 1969; *final version received July* 1971 (*Eds.*).
[2] I am very grateful to James Mirrlees for guidance and encouragement. Helpful comments were received from Avinash Dixit, a referee and an editor.
[3] This valuation function was used by Mirrlees in an unpublished paper of 1963 and by Little and Mirrlees [9].

sector and $f$ is a linear homogeneous production function. Output in the traditional sector is produced with labour alone under constant returns to scale. This is the logical counterpart of the assumption of a fixed wage to the advanced sector of the Sen-Marglin-Dixit models. This assumption is relaxed at the end of Section III. The wage to the advanced sector is $\gamma\alpha$ where $\gamma \geqq 1$ and all wages and income in the traditional sector are consumed. Thus the assumptions are as in Marglin-Dixit but we have a different instantaneous valuation function, and production and consumption assumptions in the traditional sector become explicit. There is one good. No distinction is made between public and private sector industry since the model assumes control of all the advanced sector. All advanced sector surplus is saved.

The stages which are encountered on the optimum path can be qualitatively different from those of Marglin and Dixit and the shadow wage rate determined by the model is different. This shadow wage rate throws light on the investment criteria discussion between the Polak-Buchanan-Kahn school [1] and the Dobb-Galenson-Leibenstein school. In order to understand the significance of the considerations incorporated in this model for investment criteria it is important to calculate the shadow wage rate. The optimum path has, therefore, been calculated, for two cases, which we feel illustrate the main points involved. The investment criteria aspects are indicated at the end of the paper.

## II. THE PROBLEM

An optimum path is one which maximizes the integral of the instantaneous valuation function subject to production constraints, $f(K, L)-C = \dot{K}$, $K \geqq 0$ and given $K(0)$, the wage constraint $C \geqq \gamma\alpha L$ and the employment constraint $N \geqq L \geqq 0$. We assume $r>n$ (the rate of population growth) so that existence problems are avoided [2] and we can legitimately talk of maximizing the integral. The optimum path is found using a "sufficiency lemma" proved in the appendix. The lemma is not intrinsically new [3] but is relatively easy to prove and is in a form from which it is simple to establish optimality for a fairly wide class of models (as opposed to the necessary conditions of Pontryagin *et al.* [13]).

Undefined lower-case letters are to be interpreted as the corresponding upper-case letters divided by $N$. The constraints become: $k \geqq 0$, given $k(0)$ and

$$f(k, l)-\dot{k}-nk = c \qquad \ldots(1)$$

$$c \geqq \gamma\alpha l \qquad \ldots(2)$$

$$1 \geqq l \geqq 0. \qquad \ldots(3)$$

The instantaneous valuation function becomes $lu(c/l)+(1-l)u(\alpha)e^{-\delta t}$ (where $\delta = (r-n)>0$). We could, of course, put $u(\alpha) = 0$ but it is easier to interpret the results if we carry $u(\alpha)$ through the analysis.

We now use the corollary to the sufficiency lemma (see appendix) to solve the problem. We have one state variable $k$ and two controls $c$ and $l$. We have one shadow price for the state variable (corresponding to $\boldsymbol{p}$ in the lemma) which we call $\lambda e^{-\delta t}$. It is easy to check that the problem satisfies the conditions of the corollary. Condition (*a*) gives us for $k>0$

$$\lambda f_k+\dot{\lambda}-(\delta+n)\lambda = 0. \qquad \ldots(4)$$

The Hamiltonian is $H = [lu(c/l)+(1-l)u(\alpha)+\lambda(f(k, l)-c-nk)]e^{-\delta t}$, and we must maximize $H$ with respect to $c$ and $l$. The now standard procedure of dividing $(\lambda, k)$ space

[1] For reference to, and discussion of, this argument, see Chenery [3].
[2] See Koopmans [7].
[3] See Arrow [1] and Mangasarian [10].

according to different Hamiltonian-maximizing choice of $c$ and $l$ is adopted. Equations (1) and (4) then determine, starting from any point $(\lambda, k)$, a path of $(\lambda, k)$ over time which is Hamiltonian maximizing at each point in $(\lambda, k)$ space. These paths will all satisfy conditions (*a*) and (*b*) of the corollary. Among these paths we are looking for those that satisfy also the transversality condition (*c*) which is here $\lambda k e^{-\delta t} \to 0$ as $t \to \infty$. In order to do this we sketch the stationaries of the two differential equations $\dot{\lambda} = 0$ and $\dot{k} = 0$. From these we then know the direction of the time path specified by the differential equations from any point $(\lambda, k)$. We then look for a path that tends to the intersection (if it exists) of $\dot{\lambda} = 0$ and $\dot{k} = 0$ for on this path as $t \to \infty$ we will have $\lambda k e^{-\delta t} \to 0$ as required by condition (*c*). Thus, if we can find such a path, we will have satisfied (*a*), (*b*) and (*c*) and solved our problem.

We want to maximize the Hamiltonian subject to constraints (2) and (3). We adopt a Kuhn-Tucker type procedure to do this, i.e. use shadow prices $\mu e^{-\delta t}$ and $v e^{-\delta t}$ for constraints (2) and (3) $(1 \geqq l)$. We thus look for $\mu$, $v$ so that we can find $(\bar{c}, \bar{l})$ where $(\bar{c}, \bar{l})$ maximize $\mathscr{L} = He^{\delta t} + \mu(c - \gamma\alpha l) + v(1 - l)$ and

$$\left.\begin{array}{l} \bar{c} \geqq \gamma\alpha\bar{l} \\ \mu \geqq 0 \end{array}\right\} \text{complem. and } \left.\begin{array}{r} 1 \geqq \bar{l} \\ v \geqq 0 \end{array}\right\} \text{complem.}$$

Differentiating partially with respect to $c$ we have

$$u'(c/l) - \lambda + \mu = 0 \qquad \ldots(5)$$

and with respect to $l$ we have

$$u(c/l) - (c/l)u'(c/l) - u(\alpha) + \lambda f_l - \mu\gamma\alpha - v = 0. \qquad \ldots(6)$$

Since $\mathscr{L}$ is concave in $(c, l)$ the complementary slackness conditions and (5) and (6) are sufficient to ensure maximization of $H$ subject to the constraints on $c$ and $l$.

*A*. $\mu > 0$ if and only if $\lambda > u'(\gamma\alpha)$

$\mu = 0$ if and only if $\lambda \leqq u'(\gamma\alpha)$

*B*. Suppose $\mu > 0$, i.e. (2) is binding; (2), (5) and (6) give

$$u(\gamma\alpha) - u(\alpha) + (f_l - \gamma\alpha)\lambda = v(l). \qquad \ldots(7)$$

$v$ is a decreasing function of $l$ for given $(\lambda, k)$ so $v > 0$, i.e. (3) binds at $l = 1$, if

$$u(\gamma\alpha) - u(\alpha) + (f_l(k, 1) - \gamma\alpha)\lambda > 0$$

$v = 0$ if

$$u(\gamma\alpha) - u(\alpha) + (f_l(k, 1) - \gamma\alpha)\lambda \leqq 0$$

and then $l$ is given by (7).

*C*. Suppose $\mu = 0$. We assume $u'$ is a strictly decreasing function which decreases from $+\infty$ to zero and so has an inverse defined on the positive real line; write $(u')^{-1} = g$, then (5) gives $c/l = g(\lambda)$ and (6) gives

$$v(l) = u(g(\lambda)) - \lambda g(\lambda) - u(\alpha) + \lambda f_l. \qquad \ldots(8)$$

Again, for given $(\lambda, k)$, $v$ is a decreasing function of $l$ and so (3) binds if $v(1) > 0$ and $l$ is given by (8) with $v = 0$ if $v(1) \leqq 0$.

Hence *A*, *B*, *C* define three curves which divide up $(\lambda, k)$ space as required and we call these curves separation curves *A*, *B*, *C*.

Separation curve *A*: $\lambda = u'(\gamma\alpha)$

Separation curve *B*: $u(\gamma\alpha) - u(\alpha) + (f_l(k, 1) - \gamma\alpha)\lambda = 0$

Separation curve *C*: $u(g(\lambda)) - \lambda g(\lambda) - u(\alpha) + \lambda f_l(k, 1) = 0.$

*The Canonical Policies by Region*

We call a policy that for given $(\lambda, k)$ follows the path given by the canonical differential equations, a canonical policy; the canonical differential equations are the investment equation and (4) with $(c, l)$ chosen to maximize the Hamiltonian (see appendix).

$$(a)\ \dot{k} = -nk - \gamma\alpha l + f(k, l)$$
$$\dot{\lambda} = (\delta + n - f_k(k, l))\lambda.$$

Where $l$ is chosen from

$$f_l(k, l) = \gamma\alpha - (u(\gamma\alpha) - u(\alpha))/\lambda. \quad \text{...(9)}$$

It is easy to see how to choose $l$ from (9), given $(\lambda, k)$. From $\lambda$ we read off $k/l$ from separation curve $B$ since $\eta(k/l) = \gamma\alpha - (u(\gamma\alpha) - u(\alpha)/\lambda)$, where $\eta(k) = f_l(k, 1)$. Then $l$ is just $k/(k/l)$.

$$(b)\ \dot{k} = -nk - \gamma\alpha + f(k, 1)$$
$$\dot{\lambda} = (\delta + n - f_k(k, 1))\lambda.$$

The separation curves are sketched in Figure 1. Separation curve $A$ is a horizontal straight line in $(\lambda, k)$ space. Separation curve $B$ meets $A$ at $R \equiv (u'(\gamma\alpha), k_0)$, where

$$\eta(k_0) = \gamma\alpha - (u(\gamma\alpha) - u(\alpha))/u'(\gamma\alpha),$$

if $u'(\gamma\alpha) > (u(\gamma\alpha) - u(\alpha))/\gamma\alpha$; otherwise it meets the $\lambda$ axis at $(u(\gamma\alpha) - u(\alpha))/\gamma\alpha$. From $B$'s

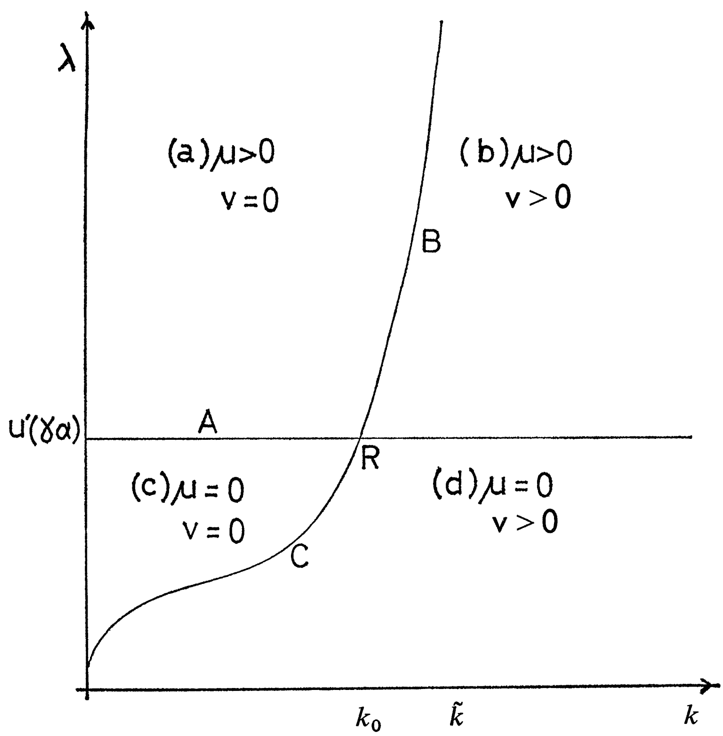


FIGURE 1

The separation curves and regions

intersection with the $\lambda$-axis or $A$, $\lambda$ increases as $k$ increases (along $B$) and $\lambda \to \infty$ as $k \to \eta^{-1}(\gamma\alpha)$. $\lambda$ increases along $C$ which terminates at $R$ (in the case where $A$ and $B$ intersect) and $C$ originates at the $\lambda$-axis where $\lambda$ is the root of $\lambda = (u(g(\lambda)) - u(\alpha))/g(\lambda)$.

The positive quadrant of $(\lambda, k)$ space is thus divided into four regions as shown in Figure 1: $\mu>0$, $v=0$ region $(a)$; $\mu>0$, $v>0$ region $(b)$; $\mu=0$, $v=0$ region $(c)$; and $\mu=0$, $v>0$ region $(d)$ (for $u'(\gamma\alpha)>(u(\gamma\alpha)-u(\alpha))/\gamma\alpha$; if the opposite inequality holds $(c)$ disappears.)

$$(c)\ \dot{k} = -nk-lg(\lambda)+f(k, l)$$

$$\dot{\lambda} = (\delta+n-f_k(k, l))\lambda \qquad \text{since } c/l = g(\lambda).$$

Where $l$ is chosen from

$$f_l = g(\lambda)-(u(g(\lambda))-u(\alpha))/\lambda. \qquad \ldots(9a)$$

Again $l$ can be readily seen from separation curve $C$.

$$(d)\ \dot{k} = -nk-g(\lambda)+f(k, 1)$$

$$\dot{\lambda} = (\delta+n-f_k(k, 1))\lambda \qquad \text{since } c = g(\lambda).$$

We must now introduce some relevant values of $k$, see Figure 2. $k_g$ is $k$ such that $f_k(k, 1) = n$, the golden rule value. I assume $f(k_g, 1)>\gamma\alpha+nk_g$. $\underline{k}<\bar{k}$ are the roots of

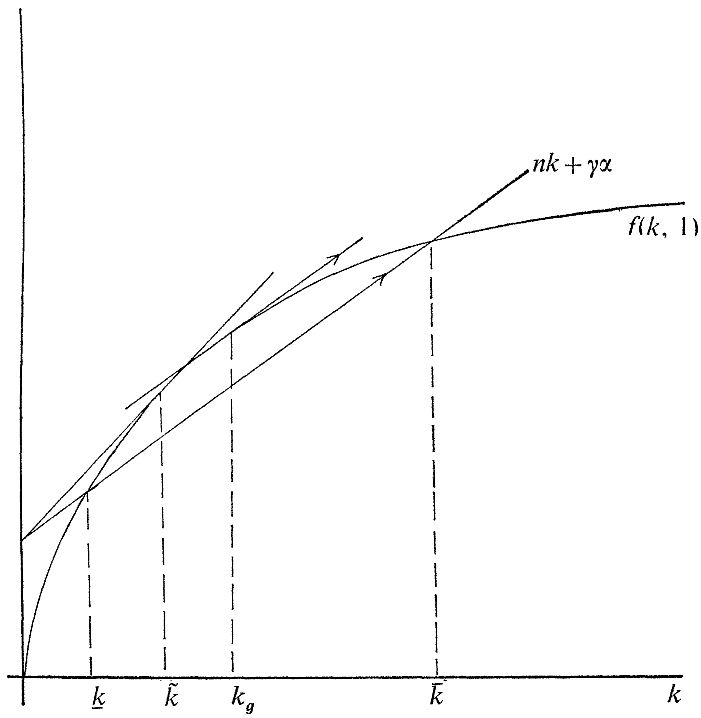


FIGURE 2

Relevant values of $k$

$f(k, 1) = \gamma\alpha+nk$, $\tilde{k}$ is $k$ such that $f_l(k, 1) = \gamma\alpha$ and $k_m$ is $k$ such that $f_k(k, 1) = \delta+n$. (This is the same notation as Dixit [3]). We know $\underline{k}<\tilde{k}<k_g<\bar{k}$ (see Fig. 2). We also know $k_m < k_g$. I assume $k_m>\underline{k}$ (so that the $\lambda$ and $k$ stationaries do intersect). Finally we have $k_0$ where $f_l(k_0, 1) = \gamma\alpha-(u(\gamma\alpha)-u(\alpha))/u'(\gamma\alpha)$. Thus $k_0 \leqq \tilde{k}$ with equality only if $\gamma = 1$. As $\gamma$ decreases to 1, $\underline{k}$ and $\tilde{k}$ decrease, $\bar{k}$ and $k_0$ increase and $k_0$ and $\tilde{k}$ tend to $\eta^{-1}(\alpha)$. For $\gamma = 1$ separation curve $B$ is vertical. We have, in principle, five cases to consider but only one needs to be examined in detail as the rest can be analysed in a similar manner.

1. $k_0 < \underline{k} < \tilde{k} < k_m$ We include here the case where $(u(\gamma\alpha) - u(\alpha))/\gamma\alpha > u'(\gamma\alpha)$ and region $(c)$ disappears.
2. $k_0 < \underline{k} < k_m < \tilde{k}$
3. $\underline{k} < k_0 < k_m < \tilde{k}$
4. $\underline{k} < k_0 < \tilde{k} < k_m$
5. $\underline{k} < k_m < k_0 < \tilde{k}$.

*The Stationaries*

The equations of the $\lambda$, $k$ stationaries are obtained by putting $\dot{\lambda}$ and $\dot{k}$ equal to zero in the canonical differential equations.

*Case* 1. It is simple to check that for this case the stationaries have the shape indicated in Figure 3.

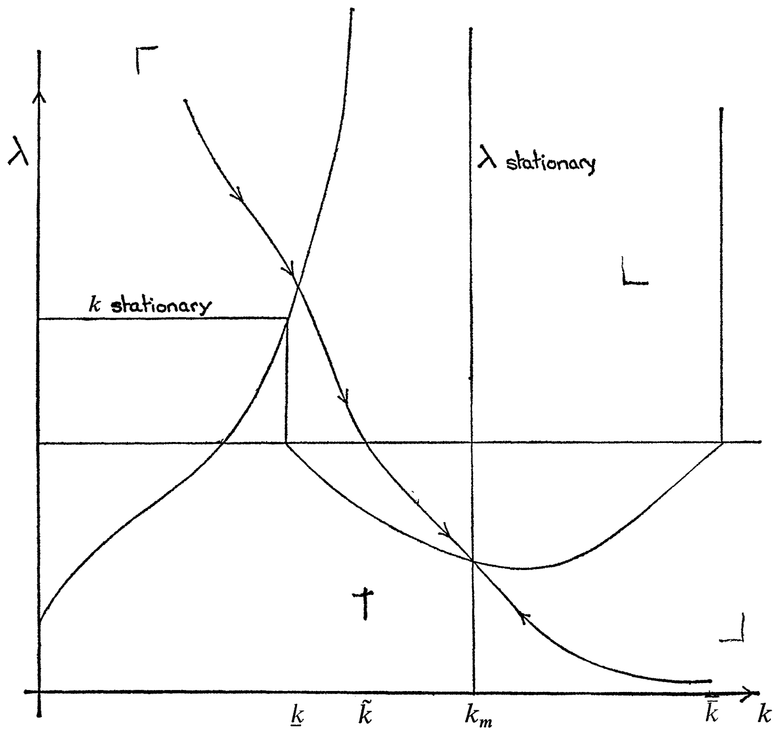


FIGURE 3

The stationaries

*Case* 2. This is like case 1 except that the $\lambda$ stationary now passes through region $(a)$. It is just $\lambda$ is a constant, $\hat{\lambda}$, such that $\eta(k_m) = \gamma\alpha - (u(\gamma\alpha) - u(\alpha))/\hat{\lambda}$, i.e. that value of $\lambda$ where $k = k_m$ meets separation curve $B$. It lies above $k = 0$ since $k_m > \underline{k}$ by assumption.

*Case* 3. This is just like case 2 except that the $k$ stationary now passes through region $(c)$, since $\underline{k} < k_0$. The $k$ stationary in region $(c)$ is $\lambda$ constant where

$$\eta(k') = g(\lambda) - (u(g(\lambda)) - u(\alpha)/\lambda)$$

and $k'$ corresponds to the point of intersection of separation curve $C$ and the $k$ stationary in ($d$).

*Case* 4. This is similar to case 3 except that the $\lambda$ stationary does not pass through region ($a$).

*Case* 5. This is similar to case 3 except that the $\lambda$ stationary now passes through region ($c$). This stationary is $\lambda$ = constant where $\eta(k_m) = g(\lambda)-(u(g(\lambda))-u(\alpha)/\lambda)$.

I assume $k_m > k'$ so that the $\lambda$ stationary lies above the $k$ stationary (and the two stationaries meet). The general picture of the stationaries is clear: we know how to draw them for regions ($b$) and ($d$). If either stationary meets separation curves $B$ or $C$ it continues in region ($a$) or ($c$) as a horizontal straight line.

*The Optimum Path*

From the phase diagrams we see that the system has a unique stationary point, $(\lambda_m, k_m)$ in each of the five cases. We must show that the stable arms through $(\lambda_m, k_m)$ extend over the whole positive $k$ axis. In cases 2, 3, 5, for $k < k_m$ the stable arm lies between the $\lambda$ and $k$ stationaries and so there exists a unique $\lambda(0)$ for each $k(0)$ such that development follows the stable arm. In cases 1 and 4 for $k < k_m$ it is easy to see that the stable arm is asymptotic to the $\lambda$ axis (otherwise $\dot{k} \to 0$ as $\lambda \to \infty$ along $k = \hat{k}$, some $\hat{k} > 0$, which we can see is impossible) and again there exists a unique $\lambda(0)$ for each $k(0)$ as required. For $k > k_m$ the stable arm also provides a unique $\lambda(0)$ for each $k(0)$ since the gradient of the stable arm is negative and we know it cannot meet the $k$-axis (since for given $k$ the optimum policies yield $\dot{\lambda} \to 0$ as $\lambda \to 0$ and $\dot{k} \to -\infty$ as $\lambda \to 0$). Hence for any $k(0) > 0$ there is a unique $\lambda(0)$ such that development follows the stable arm.

Since both $\lambda$ and $k$ tend to a constant on the stable arm this path satisfies the condition that $\lambda k e^{-\delta t} \to 0$. No other relevant path satisfies this condition since on other paths either (i) $k \to \bar{k}$ and $\lambda \sim e^{(\delta+n)t}$ or (ii) the origin is reached in finite time and the path becomes irrelevant (a path cannot reach the axes at any other point).

We have shown that for each $k(0)$ there exists a unique $\lambda(0)$ such that the three conditions of the corollary to the sufficiency lemma are satisfied. Hence for each $k(0)$ we have shown that there exists an optimum growth path. The optimum path is unique since we know from the maximum principle of Pontryagin that conditions ($a$), ($b$), ($c$) of the corollary are necessary for optimality. We are primarily interested in sufficiency, of course.

*Description of the Optimum Path*

For cases 1 and 2 we have three phases of development on the optimum path, as Dixit found for his model, when the initial $k$ is small. We begin with advanced sector wages at the minimum level and employment determined by equation (9). The optimum capital intensity in this phase is described by the shadow wage (eq. (9)). The extraction of $(\gamma\alpha)$ from the surplus is modified to take into account the consumption benefits of employment $(u(\gamma\alpha)-u(\alpha))/\lambda$. As $\lambda$ falls with accumulation the modification increases and the shadow wage and capital intensity fall. The shadow wage is discussed in section IV. Eventually we pass into region ($b$) when we have everyone employed at minimum level wages in the advanced sector and the capital intensity rises. Finally we pass into region ($d$), when we let wages rise, and follow the familiar path where the shadow price of a unit of investment per head (undiscounted) is equal to the marginal valuation of consumption per head ($\lambda = u'(c)$). An optimal path for case 1 is sketched in Figure 3. In both regions ($a$) and ($b$) the marginal valuation of a unit of advanced sector consumption per head (i.e. wages) is less than the shadow price of a unit of investment per head ($\lambda > u'(\gamma\alpha)$). In the early stages the marginal valuation of a unit of traditional sector consumption per head (i.e. peasant income) will presumably also be less than the shadow price of a unit of investment per head (early on $\lambda$ will probably be $>u'(\alpha)$). With two consumption groups there is, of course no unique price of investment in terms of consumption.

For cases 3 and 4 we pass through region ($b$) or region ($c$) but not both. Hence it might be optimum to begin raising advanced sector wages before full employment is reached.

For case 5 we have only two phases of development. We begin in region ($c$) with the marginal valuation of a unit of advanced sector wages equal to the shadow price of a unit of investment ($\lambda = u'(c/l)$) but still with population engaged in the traditional sector, i.e. we begin raising advanced sector wages from the start. Advanced sector employment is determined by equation (9$a$).

It is easy to see the economic reason why it can be optimal to raise advanced sector wages before full employment is reached; we are assuming that traditional sector output is obtained at no capital cost. It is obvious that when we do not value traditional sector output (as Marglin and Dixit do not) that we shall have the three phases of development described by Dixit, i.e. (i) unemployment and minimum wages, (ii) full employment and minimum wages, (iii) full employment and rising wages.

For suppose that, with the Marglin-Dixit valuation function it were optimum to raise wages before full employment were reached. Say that at $\hat{k}$ the optimum policy had wages higher than the minimum but also did not have full employment. For this optimum policy, therefore, the wage constraint is irrelevant and we have an answer to the familiar unconstrained optimum growth problem. However for the unconstrained problem we know that it is optimal to have full employment at any $\hat{k}$ (see e.g. Cass [2]). Hence it can never be optimal, with the Marglin-Dixit valuation function, to raise wages before full employment is reached.

Thus the explicit recognition that those not in advanced sector employment are consuming can lead to qualitatively different results and shows how the Marglin-Dixit results are obtained.

It was indicated earlier how $\underline{k}$, $\bar{k}$, $k_0$ and $\tilde{k}$ change with $\gamma$. We can make $k_m$ as small as we please by increasing $r$ the rate at which valuations are discounted. High $\gamma$ and low $\delta$ mean that the advanced sector minimum wage is much higher than peasant income and that we have a high long-run target for capital per head (high $k_m$); this gives case 1. Low $\gamma$ and high $\delta$ means that advanced sector minimum wages are close to peasant income and we do not have ambitious long-run targets for the capital stock; in these circumstances case 5 might arise. Provided $k_m < \eta^{-1}(\alpha)$ there will be a $\gamma > 1$ such that case 5 arises.[1]

## III. CALCULATION OF PATH AND SHADOW WAGE

In this section we show how the optimum growth path can be calculated. We have to solve the canonical differential equations for each of the relevant regions through which it passes. For most cases the three relevant regions will be ($a$), ($b$), ($d$). The canonical differential equations give us a first order differential equation between $\lambda$ and $k$ for the optimum path. In order to solve this first order differential equation uniquely we need to have a constant term, i.e. for one $k$ we must know $\lambda$. In our situation therefore we must begin solving the equations in region ($d$) since only here do we know the optimum $\lambda$ for a given $k$, i.e. $\lambda_m$ for $k_m$. The optimal path in region ($d$) is found by a technique due to Mirrlees [12]. We solve the differential equation using an infinite series expansion around ($\lambda_m$, $k_m$). The optimal path for example (i) is drawn in Figure 4. We trace this path back to where it meets separation curve $A$, at $P$ say (see Fig. 4). From $P$ we must use the canonical differential equations corresponding to region ($b$). For the two cases presented here it was possible to solve analytically the corresponding differential equation in $\lambda$

[1] The two calculations presented here based on judgments of $r$ and $u$ suggest that only case 1 will be relevant in practice.

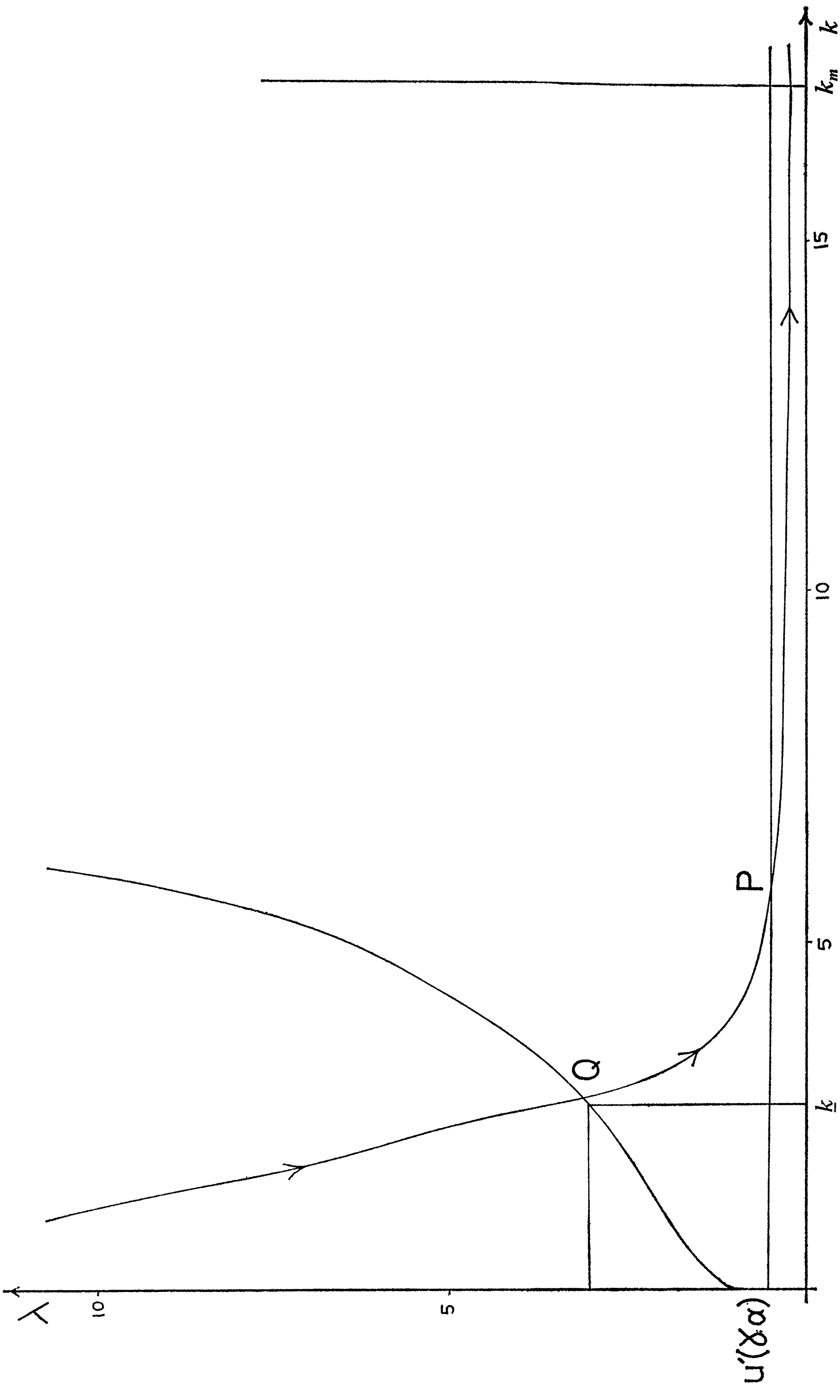


FIGURE 4
The calculated optimum growth path (for example (i) § III)

and $k$. This path is then traced back to where it meets separation curve $B$, at $Q$ say. It was again possible to solve analytically the differential equation appropriate for region ($a$). This gives us the whole relevant optimum path. We give the parameters used and the results obtained for two examples.

*Example*[1] (*i*)

Production is Cobb-Douglas $f(k, l) = Ak^bl^{1-b}$ where $b = \frac{1}{3}$ and units are chosen so that $A = 1$. Suppose initially; $\frac{1}{5}$ of the population is in the advanced sector and $\frac{4}{5}$ in the traditional sector; the capital output ratio is 4 and that each sector produces $\frac{1}{2}$ of *GNP*. These assumptions imply that initially traditional sector income per head is $\frac{5}{8}$ of *GNP* per head, $k(0) = \frac{8}{5}$ and initial *GNP* per head is $\frac{4}{5}$ and that $\alpha = \frac{1}{2}$. We assume that the wage equals the marginal product in the advanced sector so that the wage bill in the advanced sector is $\frac{2}{3}$ output and $\gamma = \frac{8}{3}$. The valuation function assumed is $u(c) = -1/c$. This choice of valuation function means that we value an extra unit of consumption given to group $A$ four times as much as a unit given to group $B$ if group $B$ has consumption per head twice as large as group $A$. Population growth is at 2 per cent and instantaneous valuations are discounted at 5 per cent. These are all the parameters we need to calculate the optimum growth path.

It is easy to check that $\underline{k} = 2{\cdot}67$, $\tilde{k} = 8$ and $k_m = 17{\cdot}2$. We thus have an example of case 1. Working back from $(\lambda_m, k_m)$ gives the point of intersection of the path with curve $A$ at $P$ where $\lambda = \frac{9}{16}$, $k = 5{\cdot}80$. The path meets $B$ at $Q$ where $k = 2{\cdot}89$, $\lambda = 3{\cdot}26$. Finally working back to $k(0)$ we have $\lambda(0) = 8{\cdot}25$ at $k(0) = 1{\cdot}60$.

We can now describe the optimum development path. We can also calculate the time of the various phases and the shadow wage: $\gamma\alpha-(u(\gamma\alpha)-u(\alpha))/\lambda$ (see section IV). We begin with capital per head of population at 1·6. We employ labour until its marginal product is equal to the shadow wage which is 88·6 per cent of the market wage. Full employment is reached after 12·50 years when the shadow wage has fallen to 55·9 per cent of the market wage and capital per head of population is 2·9. For the next 19·70 years wages are kept at the minimum level until capital per head rises to 5·8. Thereafter consumption and investment are kept equally valuable and we follow the familiar optimum growth path with capital per head and wages increasing.

*Example*[2] (*ii*)

Production is Cobb-Douglas with $b = \frac{1}{2}$. Initially the advanced sector capital-output ratio is 3 and $\frac{1}{3}$ *GNP* produced in the traditional sector. Population growth is 3 per cent per annum, instantaneous valuations are discounted at 4 per cent and the valuation function is $u(c) = \log_e c$. Other assumptions are as in example (i). These assumptions imply $\gamma = 4$, $\underline{k} = 2{\cdot}48$, $\tilde{k} = 9$, $k_m = 156{\cdot}25$. We thus have an example of case 1. The optimum development path is as follows. We begin with capital per head of population at 1·8. We employ labour until its marginal product is 85·7 per cent of the market wage. Full employment is reached after 7·48 years when the shadow wage has fallen to 40·5 per cent of the market wage and capital per head is 3·53. For the next 9·79 years wages are kept at their minimum level until capital per head rises to 8·9. Thereafter the familiar optimum growth path, with consumption and investment equally valuable, is followed.

These results should not be taken too literally but some lessons can be learnt. The main conclusion is that the initial shadow wage is high in relation to the market wage; 88·6 per cent of the market wage in the first example and 85·7 per cent in the second example. The reason is the high initial value of $\lambda$: 8·25 (in example (i)) compared with $u'(\gamma\alpha)$ of $\frac{9}{16}$ and $u'(\alpha)$ of 4. Thus the value of investment in terms of a marginal increment to con-

[1] The figures used here are based roughly on the Indian economy. The outcome of the calculation should not be taken too seriously.

[2] Kenyan data was used here. The qualification of the previous footnote applies.

sumption at the level of advanced sector wages is 14·7 and in terms of a marginal increment to consumption at the level of traditional consumption is 2·1. In example (ii) the corresponding values are $\lambda(0) = 6{\cdot}45$, $u'(\gamma\alpha) = \frac{2}{3}$ and $u'(\alpha) = \frac{8}{3}$. It is probably the case that further examples of this kind of analysis will also have high initial values of the shadow wage. The reason is that the valuation function used imputes high values to investment relative to consumption at current urban wage rates; in other words the shadow price of the wage constraint rises very quickly as the constraint bites (see Fig. 4 and equation (5)).

In this model the shadow wage falls during the initial stages (see the discussion of the optimum path). The reason is that the accumulation of capital results in the wage constraint biting less strongly so that $\lambda$ and $\mu$ both fall (see equation (5)). In a model with a more thorough articulation of the traditional sector or with different wage and migration assumptions this is less likely to occur. The assumption of constant returns in the traditional sector is also responsible for the rapid achievement of full employment—$12\frac{1}{2}$ years in the first example and $7\frac{1}{2}$ years in the second. Clearly no one will believe this is possible in practice and introducing diminishing returns with the traditional sector would certainly extend this time estimate, although the kind of flexibility assumed in this model will still lead to underestimates of the time. The assumption that all profits are saved biases the result in the same direction.

The main reason for the assumption of constant returns in the traditional sector was technical. For the diagrammatic technique of analysis used here it is necessary to remove the explicit time dependence from the problem (apart from the $e^{-\delta t}$ term in the maximand). This causes difficulties with the $u(\alpha)$ term in the maximand: we must make an assumption that ensures $\alpha$ is a function of $l$ only. There are at least three ways of doing this. First, the constant returns assumption used here. Secondly, we can assume population is constant.[1] Thirdly, we can assume technical progress is of just such a kind that $\alpha$ is a function of the proportion, $l$, of the population in the traditional sector. Clearly none of these assumptions is entirely plausible. The assumption used here was chosen since it is the simplest approach that allows for population growth and brings out the main points obtainable from this kind of analysis. The two main consequences for the optimum development path have been indicated above. We give (9)′, the modification of equation (9) when we allow for diminishing returns, to help the discussion of the shadow wage in the next section. It can easily be checked by the reader (compare also with the appendix to [8]).

$$_{l} = \gamma\alpha - (u(\gamma\alpha) - u(\alpha))/\lambda - u'(\alpha)(\alpha - m)/\lambda + \frac{\gamma l}{1-l}(\alpha - m)(1 - u'(\gamma\alpha)/\lambda) \qquad \ldots(9)'$$

## IV. CONCLUSIONS

The first conclusions concern investment criteria and in particular the shadow wage (i.e. the value to which the marginal product of labour should be equated) which is given by (9). Discussion [2] of the appropriate shadow wage, $w^*$, in underdeveloped countries has mainly been between the social marginal productivity school (who claim $w^*$ is the marginal product in the traditional sector) and the marginal reinvestment school (who claim $w^*$ is equal to the actual wage, in order to maximize surplus and investment). Sen [14] has shown how the schools can be distinguished by their different view of the shadow price, $\lambda^*$, of investment in terms of consumption; the former regard $\lambda^*$ as 1, the latter as infinite. This however is not the whole story. Our view of $w^*$ should not only be affected by our views concerning the relative value of consumption and investment (i.e. of consumption now or later) but also our views concerning the distribution of consumption

[1] This problem is tackled in the appendix to Ch. 2 of the author's D.Phil. thesis, Oxford University, 1971.

[2] See Chenery [3].

M

now. Thus equation (9) counts the benefits of the transfer of a worker from a lower to a higher income class [1] (deflated by the shadow price of investment). However, if we put $\gamma = 1$ in (9) we have $w^* = \gamma\alpha$ the market wage, i.e. we require the surplus-maximizing shadow wage. The economic reason is clear: if everyone will be consuming the same whether in the advanced or traditional sector *regardless* of our employment policy then we serve the future best by growing as fast as possible. This possibility could not arise in the Dixit model since his shadow wage, $w_d$, is always less than the market wage:

$$w_d = \gamma\alpha - (\gamma\alpha u'(\gamma\alpha l)/\lambda)$$

(my notation for eq. (6) in Dixit [4]). Our shadow wage is modified slightly if we allow diminishing returns in the traditional sector: the first extra term in (9)′ (as compared with (9)) indicates the benefit of the increased consumption in the rural sector (the excess of the marginal worker's consumption over production) and the second the cost of pushing up the institutional wage. If $\lambda$ is large the second extra term dominates the first and $w^*$ may even be higher than the market wage. It should be noted that the difference between the market wage and the shadow wage is *not* necessarily the appropriate wage subsidy. The reason is that a subsidy is a transfer out of government funds.[2]

This model has focused on just two aspects of the shadow wage; the shadow price of investment and the examination of the degree to which allowance should be made for the consumption benefits of transferring a worker from a lower to a higher income class. These consumption benefits *should* be counted but with a shadow price of investment as high as that implied by the two calculations this allowance is small. There are, of course, other important aspects of the shadow wage and the construction of simple models has also thrown light on these; see e.g. Stiglitz [15] and Hornby [5] who discuss the problems of urban-unemployment and the price of food respectively. These analyses also point to high shadow wages, given high institutional wages.

We conclude with a brief summary. The optimal development path has been characterised in a model similar to that of Marglin and Dixit. Calculations of the path have been presented in two simple cases in order to throw light on the magnitudes involved. The initial shadow price of investment was found to be high in both cases so that the shadow wage was close to the market wage. The important difference from the Marglin-Dixit model is the different valuation function. This yields the possibility of a different time-sequence of stages of development from that found by Dixit. The other important difference is the possibility that the shadow wage may be equal to or higher than the market wage; this cannot arise in the Dixit formulation.

## APPENDIX: THE SUFFICIENCY LEMMA

Suppose we have state variables $\boldsymbol{x}(t)$ and control variables $\boldsymbol{y}(t)$. We denote $(\boldsymbol{x}(t), \boldsymbol{y}(t))$ by $z_t$ which is contained in some set of definition $Z$ (which may depend on $t$). $v_t$ is an instantaneous valuation function defined on $Z$. We say a programme $(z_t)$ is at least as good as $(z_t')$ if given $\varepsilon > 0$ there exists a $T_0$ such that for all $T > T_0$,

$$\int_0^T v_t(z_t)dt \geqq \int_0^T v_t(z_t')dt - \varepsilon.$$

A programme is optimal if it satisfies all the constraints and is at least as good as any other programme satisfying the constraints. (This criterion is slightly weaker than the overtaking criterion used in e.g. Mirrlees [12]). The constraints are $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{y}, t)$, a given $\boldsymbol{x}(0)$ and $(\boldsymbol{x}, \boldsymbol{y})$ contained in some set $W$ (which is a subset of $\{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \geqq \boldsymbol{0}\}$ and which may depend on $t$). A programme satisfying the constraints is called feasible.

1 Similar results are obtained in the appendix to [9].
2 This point is elaborated in Ch. 4 of the author's D.Phil. thesis, Oxford University, 1971.

**Lemma.** *If there exist* $\mathbf{p}(t)$ *(non-negative and piece-wise differentiable) and* $(z_t^*)$ *(feasible) such that* (i) $z_t^*$ *maximizes* $H(z_t, \mathbf{p}, t)+\dot{\mathbf{p}}.\mathbf{x}$ *each* $t$, *subject to* $\mathbf{x}(0)$ *given and* $z_t \in W$ *(apart from the, at most countable number of, points where* $\dot{\mathbf{p}}$ *is undefined) and* (ii) $\mathbf{p}.\mathbf{x}^* \to 0$ *as* $t \to \infty$; *where* $H(z_t, \mathbf{p}, t) = v_t(z_t)+\mathbf{p}.f(\mathbf{x}, \mathbf{y}, t)$, *then* $(z_t^*)$ *is an optimal programme.*

*Proof.* Consider any other programme $(z_t')$ satisfying the constraints

$$\int_0^T (H(z_t^*, \mathbf{p}, t)+\dot{\mathbf{p}}.\mathbf{x}^*)dt \geqq \int_0^T (H(z_t', \mathbf{p}, t)+\dot{\mathbf{p}}.\mathbf{x}')dt$$

for all $T$ by condition (i)

$$\int_0^T (H(z_t^*, \mathbf{p}, t)-\mathbf{p}.\dot{\mathbf{x}}^*)dt+[\mathbf{p}.\mathbf{x}^*]_0^T \geqq \int_0^T H(z_t', \mathbf{p}, t)-\mathbf{p}.\dot{\mathbf{x}}')dt+[\mathbf{p}.x']_0^T$$

$$\int_0^T v_t(z_t^*)dt+\mathbf{p}(T).\mathbf{x}^*(T) \geqq \int_0^T v_t(z_t')dt+\mathbf{p}(T).x'(T)$$

$$\int_0^T v_t(z_t^*)dt \geqq \int_0^T v_t(z_t')dt-\mathbf{p}(T).\mathbf{x}^*(T).$$

Thus, by condition (ii), $(z_t^*)$ is at least as good as any other feasible programme and is optimal.

**Corollary.** *Suppose* $H(z_t, \mathbf{p}, t)$ *is concave and differentiable in* $z_t$, $\exists$ *an* $\varepsilon$ *s.t.*

$$(\mathbf{x}^*+\mathbf{h}, \mathbf{y}^*) \in W \ \forall \ \|\mathbf{h}\| < \varepsilon,$$

$\mathbf{p}$ *is as in the lemma, and* $W$ *is convex.*

(*a*) $\dfrac{\partial H}{\partial x_i}(\mathbf{x}^*, \mathbf{y}^*, \mathbf{p}, t)+\dot{p}_i = 0$ *each* $i$[1]

(*b*) $\mathbf{y}_*$ *maximises* $H(\mathbf{x}^*, \mathbf{y}, \mathbf{p}, t)$ *subject to* $(\mathbf{x}^*, \mathbf{y}) \in W$

(*c*) $\mathbf{p}.\mathbf{x}^* \to 0$ *as* $t \to \infty$

*then* $(\mathbf{x}^*, \mathbf{y}^*)$ *is an optimum programme.*

*Proof.* (*a*) and (*b*) imply (i) with the stated assumptions.

The lemma is slightly more general than Proposition 5 in Arrow [1]. Arrow's conditions (the main one being that $\max_{\mathbf{y}} H(\mathbf{x}, \mathbf{y}, \mathbf{p}, t)$ is concave in $\mathbf{x}$) imply that

$$H(\mathbf{x}, \mathbf{y}, \mathbf{p}, t)+\dot{\mathbf{p}}.\mathbf{x}$$

is maximized at $(\mathbf{x}^*, \mathbf{y}^*)$ so that the conditions of the lemma are satisfied. The lemma can, however, be used in the presence of non-concavities provided that we can check that $H+\dot{\mathbf{p}}.\mathbf{x}$ is maximized. $H+\dot{\mathbf{p}}.\mathbf{x}$ can be interpreted as national income including capital gains. The corollary is weaker than Arrow's theorem but is in a more useful form—it is not always easy to check directly the concavity of $\max_{\mathbf{y}} H(\mathbf{x}, \mathbf{y}, \mathbf{p}, t)$. It is strong enough to cover many economic problems—including the one analysed in the text. The corollary has also been noted by Kamien and Schwartz [6]. The lemma, the corollary and Arrow's theorem are different versions of the extension of Mangasarian [10] to the case of an infinite time horizon. The integrating-by-parts argument is also to be found in Mirrlees [12].

1 We have assumed that constraints do not "bind in the $x$ direction". If they do we can include Lagrange multipliers.

## REFERENCES


[1] Arrow, K. J. " Applications of Control Theory to Economic Growth ". Lectures in Applied Mathematics, No. 12. *Mathematics of the Decision Sciences Pt. II* (1968) pp. 85-119.

[2] Cass, D. " Optimal Growth in an Aggregative Model of Capital Accumulation ", *Review of Economic Studies*, **32** (1965).

[3] Chenery, H. B. " Comparative Advantage and Development Policy ", *American Economic Review*, **51** (1961).

[4] Dixit, A. K. " Optimal Development in the Labour-Surplus Economy ", *Review of Economic Studies*, **35** (1968).

[5] Hornby, J. M. " Investment and trade policy in the dual economy ", *Economic Journal*, **78** (1968).

[6] Kamien, M. I. and Schwartz, N. L. " Sufficient Conditions in Optimal Control Theory ", *Journal of Economic Theory*, **3** (1971).

[7] Koopmans, T. J. " On the Concept of Optimal Growth ", in *The Econometric Approach to Development Planning* (Pontificia Academia Scienterum, Vatican, 1965).

[8] Lewis, W. A. " Economic Development with Unlimited Supplies of Labour ", *The Manchester School*, **22** (1954)

[9] Little, I. M. D. and Mirrlees, J. A. *Manual of Industrial Project Analysis in Developing Countries. Volume II: Social Cost Benefit Analysis*, OECD (1969).

[10] Mangasarian, L. " Sufficient Conditions for the Optimal Control of Non-linear Systems ", *SIAM*, *J. Control*, **4** (1966).

[11] Marglin, S. A. " Industrial Development in the Labour-surplus Economy " (Harvard, 1966, unpublished).

[12] Mirrlees, J. A. " Optimum Growth when Technology is Changing ", *Review of Economic Studies*, **34** (1967).

[13] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V. and Mischenko, E. F. *The Mathematical Theory of Optimal Processes* (Interscience, New York, 1962).

[14] Sen, A. K. *Choice of Techniques*, 3rd ed., especially new introduction (Basil Blackwell, 1968).

[15] Stiglitz, J. E. " Rural-Urban Migration, Surplus Labour, and the Relationship between Urban and Rural Wages ", *Eastern Africa Economic Review* (December 1969).