Cowles Foundation Paper 269

CHAPTER 5

*Intertemporal Distribution and 'Optimal' Aggregate Economic Growth**

TJALLING C. KOOPMANS

"... it is assumed that we do not discount later enjoyments in comparison with earlier ones, a practice which is ethically indefensible and arises merely from the weakness of the imagination, . . ."

F. P. Ramsey [1928].

"... we feel less concerned about future sensations of joy and sorrow simply because they do lie in the future. Consequently we accord to goods which are intended to serve future ends a value which falls short of the true intensity of their future marginal utility."

E. von Böhm-Bawerk [1921, II, p. 268].

"On the assumption ... that a government is capable of planning what is best for its subjects, it will pay no attention to pure time preference, a polite expression for rapacity and the conquest of reason by passion."

R. F. Harrod [1948, p. 40].

"In such an ideal loan market, therefore, where every individual could freely borrow or lend, the rates of preference or impatience for present over future income for all the different individuals would become, at the margin, exactly equal to each other and to the rate of interest."

Irving Fisher [1930, p. 106].

"Most people are of the humour of an old fellow of a college, who, when he was pressed by the Society to come into something that might redound to the good of their successors, grew very peevish; 'We are always doing,' says he, 'something for posterity, but I would fain see posterity do something for us,' "

Joseph Addison, *The Spectator*, Vol. VIII, No. 583, August 20, 1714.

---

* This paper has resulted from research under a grant from the National Science Foundation. An earlier version was presented at a joint meeting of the Econometric Society and the American Economic Association held at Boston, December, 1963. I am indebted to James Tobin for highly valuable comments, and to Koen Suryatmodjo for very fine draftsmanship.

Reprint From
Ten Economic Studies in the
Tradition of Irving Fisher
Published by John Wiley & Sons, Inc., 1967

Scratch an economist and you find a moralist underneath. The clearest exception to this rule for once truly proves the rule: Some of our most illustrious British colleagues have cast all dissimulation aside. No scratching is needed in their case!

It is true that, in the quotations given, Ramsey and Harrod were commenting on possible time preference underlying governmental planning. In contrast, Böhm-Bawerk and Irving Fisher were concerned, in a more detached manner, with the observable time preferences of individuals and with analyzing the market effects of these preferences. However, the context makes clear that Harrod has little use for a positive "pure time preference" under any circumstances:

"Time preference in this sense is a human infirmity, probably stronger in primitive than in civilized man."*

Moreover, in all societies where, in one way or another, individual wants and desires do have an effect on government action, individual and social time preferences are inevitably connected. So we do have an ethical problem here, either at the individual level, or in explicit regard to planning.

What is at issue is clearly an intertemporal distribution problem: that of balancing the consumption levels of successive generations, and of successive stages in the life-cycle of a given cohort of contemporaries. The most pertinent decisions—individual, corporate, or governmental—are those that determine investment in physical capital, in human capital, and in research and development. Investments in physical capital, if well made, augment future consumption through an increase in future capital-labor ratios. Investment in human capital raises the quality of labor and, one hopes, of life. Successful research and development augment future output from given future capital and labor inputs through the development of better techniques of production.

Recent research on models of optimal growth has clarified the boundaries within which there is scope for ethical judgment regarding time preference.† The purpose of this chapter is to make the preoccupations and some of the findings of these researches plausible

* Harrod [1948], p. 37.

† See Koopmans [1965, 1967] and papers by Cass, Inagaki, Malinvaud, Mirrlees, Phelps, Samuelson, von Weizsäcker, there cited.

to a larger readership through a diagrammatic analysis of one particular model of "optimal" growth that is highly stylized and simplified, yet representative of more realistic models in regard to the particular question at issue. We shall concentrate on exposition rather than evaluation of the findings. For some evaluative remarks, and a survey of a wider range of findings, see Koopmans [1967].

We shall make no assumptions about the particular institutional form of the economy discussed. The simplest interpretation is in terms of an economy in which growth rates are centrally planned in a manner capable of implementation. It is hoped that the analysis can also serve as background for the discussion of growth policies in an individual or corporate enterprise society, or under conditions of less perfect and dependable planning. In either case, the main aim is to obtain insight into the effect of, and the scope for, time preference.

1. Assumptions Regarding Production and Population Growth

Our model has a single good, capable of serving as consumption good or as capital good, as desired. The net excess of its output flow over its consumption flow automatically becomes a net addition to the capital stock, which in turn affects output from a given labor input. Technology and the quality of labor are constant over time. Hence only the first of the three types of investment decisions mentioned arises in the model.

Technology is represented by a production function $F(L, K)$ giving the rate of output as a function of the labor force L and the capital stock K. This function, defined and assumed twice differentiable for all nonnegative L, K, has the following further properties:

(*a*) $F(L, 0) = 0$ (no capital no output)
(*b*) $F'_K(L, K) > 0$ for all $L > 0, K \geq 0$ (the marginal productivity of capital is positive for all factor combinations with some labor)
(*c*) $F''_K(L, K) < 0$ for all $L > 0, K \geq 0$ (the marginal productivity of capital decreases as capital is increased while labor is held constant)
(*d*) $F(0, K) = 0$, (no labor no output)
(*e*) $F(L, K) = LF(1, K/L) = Lf(K/L)$ (that is, constant returns to scale)

The popularity of assumption (e) is due more to the analytical simplifications it permits than to its claim to realism. In the present case, (e) opens our problem up for the use of diagrams on the printed page. It allows the production function of two variables, L and K, to be derived from the per-worker production function $f(k)$ that depends only on the single variable $k = K/L$, capital per worker. To prepare for these diagrams, we translate the assumptions (a) through (d) in terms of that function $f(k)$. Using (e), we derive* the following from (a), (b), (c):

$(a')\ f(0) = 0, \quad (b')\ f'(k) > 0, \quad (c')\ f''(k) < 0.$

The per-worker production function $f(k)$ therefore has a form as indicated in Figure 1. Beginning at $f(0) = 0$, it rises for all $k \geq 0$, but at a decreasing rate as k increases.

We did not specify counterparts to (b) and (c) that refer to increases in labor instead of in capital, because these counterparts are implied in (b), (c), and (e)—which shows the force of (e). However, (d) gives us new information about $f(k)$,

$(d')\ \lim_{k\to\infty} \frac{f(k)}{k} = 0$ (the average product of capital tends to zero as the capital per worker is increased indefinitely†)

Geometrically (see Figure 1), any rising straight line $y = \lambda k$, $\lambda > 0$, through the origin will eventually cross the curve $y = f(k)$, as k is made larger and larger—no matter how small the slope λ.

We shall assume exogenously given exponential labor force growth

$$L_t = e^{\lambda t}, \qquad \lambda > 0, \tag{1}$$

choosing the initial labor force at time $t = 0$ as the unit of labor force, $L_0 = 1$. We shall speak as if the labor force is the entire population, merely to avoid the extra symbol that would be required if we assumed the labor force to be a constant fraction of the population.

Using a continuous time variable t, and using dotted symbols for time derivatives such as $\dot{K}_t = dK/dt$, output is allocated to consumption C_t and to net capital formation $\dot{K}_t$ according to the identity

$$F(L_t, K_t) = C_t + \dot{K}_t. \tag{2}$$

* Since $F'_K = f'(K/L)$, $F''_K = (1/L)f''(K/L)$.

† *Proof*: $0 = F(0, 1) = \lim_{L\to 0} F(L, 1) = \lim_{L\to 0} Lf\left(\frac{1}{L}\right) = \lim_{k\to\infty} \frac{f(k)}{k}$, taking $L = 1/k$.

The corresponding identity in terms of per-worker capital k_t and consumption $c_t = C_t/L_t$ is obtained* by dividing through by L_t,

$$f(k_t) = c_t + \lambda k_t + \dot{k}_t. \tag{3}$$

This identity, basic in all that follows, says that per-worker output is allocated to three ends: (1) per-worker consumption c_t, (2) an investment of λk_t needed merely to keep the per-worker capital stock constant (that is, to keep the absolute capital stock K_t growing in proportion to the labor force), and (3) a net rate of increase $\dot{k}_t$(positive or negative) in the capital stock per-worker.

Equation (3) assumes that capital does not depreciate. A simple reinterpretation will cover the case of exponential depreciation at a rate δ as well: One replaces λ in (3) by

$$\lambda^* = \lambda + \delta. \tag{4}$$

We will not pursue this reinterpretation here.

2. The Golden Rule of Accumulation

Before discussing the choice of the objective of growth policy in general, we look at a special problem so defined as to leave only one obvious choice of the objective.

Suppose that the economy of the island Roswesri Adelphi satisfies all the assumptions we have made. Upon its admission to the United Nations, the World Bank offers, as a once-and-for-all gift, to supply whatever additional "capital" is needed to bring the total capital stock at $t = 0$ to any level the newly sovereign government specifies. This generous offer is subject to only one condition, deemed indefinitely enforceable by all concerned: the people of Roswesri Adelphi must at all times $t \geq 0$ allocate just enough of their output to investment to keep the per-worker capital stock constant,

$$k_t = k \qquad \text{for all } t \geq 0. \tag{5}$$

What initial capital-per-worker k should the government ask for? Inserting (5) in equation (3) shows that consumption per worker

$$c_t = c = f(k) - \lambda k \tag{6}$$

---

* Because $\dot{k}_t = (d/dt)(K_t e^{-\lambda t}) = (\dot{K}_t - \lambda K_t)e^{-\lambda t} = \dot{K}_t/L_t - \lambda k_t$.

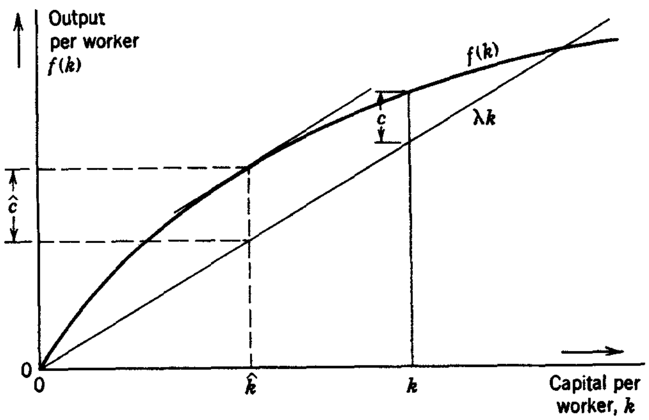

Figure 1 The golden rule of accumulation.

has also become a constant. Figure 1 shows the construction of c as the excess, at the point k, of the curve $f(k)$ over the straight line λk. Obviously, the only sensible objective is to maximize c, once-and-for-all. This requires (Figure 1) choosing that $k = \hat{k}$ for which the slope $f'(k)$ of the tangent to the curve $f(k)$ equals λ,

$$f'(\hat{k}) = \lambda, \qquad \hat{c} = f(\hat{k}) - \lambda\hat{k}. \tag{7}$$

Among all paths with a constant per-worker capital stock, the highest consumption per worker is attained and maintained by that path on which the marginal productivity of capital equals the (constant) rate of population growth.

This simple but important proposition was discovered many times over in the late 'fifties and early 'sixties. Nine independent discoverers are listed by one of them, Phelps [1966, pp. 3, 4], in what is by now the fullest discussion of its many ramifications. The nine papers vary in the generality of their assumptions. Some of them permit labor-augmenting technical progress. The policy of maintaining the per-worker capital stock that, once attained, permits the highest consumption per worker has been called the *golden rule of accumulation* by Phelps, because then

. . . each generation saves (for future generations) that fraction of income which it would have past generations save for it . . .*

* [1966], p. 5.

The path resulting from the policy has been called the *golden rule path*.

If, by way of comparative dynamics, one considers an archipelago with different population growth rates on different islands, then as λ approaches zero the capital-per-worker $\hat{k}$ prescribed by the golden rule approaches the unattainable infinity—unless one reintroduces a positive rate of depreciation.

3. Choice of the Objective

The golden rule path is, of course, available only after the required initial capital stock has been attained. For any different, historically given, initial capital stock one needs a more discriminating criterion. But even if the requisite per-worker capital stock were to be on hand, we must remember that the rule was derived from an arbitrary condition of the unchangeability of that capital-labor ratio. We must still explore what an economy not bound by such a condition might want to do.

We shall first discuss this problem for a constant population. The criterion most used is the sum over time (literally a sum for discrete time, an integral for continuous time) of future utilities discounted to the present time. One postulates a *utility function* $u(c)$ that expresses the utility flow generated at any time in the future at which consumption flows at the positive rate c. The function is assumed to increase with c, but at a decreasing rate,

$$u'(c) > 0, \qquad u''(c) < 0 \qquad \text{for all } c > 0. \tag{8}$$

Finally, to avoid the possibility that a zero rate of consumption could temporarily be optimal, we give the utility curve a vertical tangent at $c = 0$,

$$\lim_{c \to 0} u'(c) = \infty. \tag{9}$$

Figure 2 indicates a possible form of $u(c)$, with $u(0)$ finite. Another form, with $\lim_{c \to \infty} u(c) = -\infty$, can be seen in Figures 5 to 8.

As the objective of growth policy we now consider a *utility functional* that depends on an entire consumption path c_t, $0 \leq t \leq T$, in the form of an integral

$$U_T = \int_0^T e^{-\rho t} u(c_t)\, dt, \qquad 0 < \rho < 1. \tag{10}$$

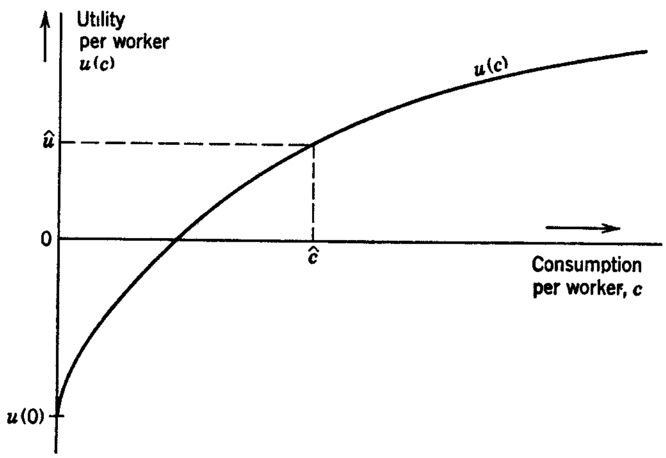

Figure 2 The utility function.

ρ is the constant (instantaneous) *discount rate*, $e^{-\rho}$ the *discount factor* for one unit of time (one year, say).

Note that if we choose to make a linear change

$$v(c) = \alpha u(c) + \beta, \qquad \alpha > 0, \tag{11}$$

in the utility scale, similar to the change from Fahrenheit to centigrade in the measurement of temperature, then the utility functional is rescaled in the same way,

$$V_T = \int_0^T e^{-\rho t} v(c_t)\, dt = \alpha U_T + \gamma.$$

Therefore a path optimal with reference to the u-scale remains optimal if the v-scale is used instead.

If c_t is a continuous consumption path, the quantity

$$\phi = e^{\rho} \cdot \frac{u'(c_0)}{u'(c_1)} \tag{12}$$

is the ratio of the present marginal utility of one small extra unit of consumption *now* to the *present* marginal utility of the sure prospect of an extra unit of consumption *one year from now*. Its excess over unity, $\phi - 1$, represents what Irving Fisher [1930, Ch. IV] has called "time preference" or, synonymously in his usage, "impatience." If consumption is the same at the two points in time, $c_0 = c_1$, then $\phi = e^\rho$ is the reciprocal of the discount factor, and $\phi - 1$ is Harrod's "pure time preference," of which he disapproves. If, on the other hand, $\rho = 0$, so pure time preference is absent, but $c_0 \neq c_1$, then Fisher's impatience $\phi - 1$ arises solely from the fact that the higher rate of consumption entails the smaller marginal utility. In this connection, Harrod argues persuasively that a society anticipating rising consumption would exhibit a positive interest rate even in the absence of pure time preference.

Since the second factor in (12) is a *ratio* of *marginal* utilities, time preference is similarly unaffected by any linear scale change (11). However, a *nonlinear* scale change would affect time preference as previously defined. This did not worry either Fisher or Harrod, since both attribute a natural cardinal meaning to utility. If one is reluctant to do so, as I am, then one must fall back on the statement that there is one special class of scales, all linearly related, in which the utility functional (10) has the simple form of an integral over discounted "utilities" (so scaled). If thereafter one uses the expression marginal utility, it is to be tacitly understood that one uses the term "utility" with reference to a scale of that special class. Although the use of such a scale is not obligatory, it brings the benefits of postulated simplicity.

This point deserves emphasis because the simplicity of (10), and with it the formal cardinality of the utility scale, are bought at the price of an implication of noncomplementarity of consumption levels at different points in time.* That is indeed a steep price!

In maximizing the integral (10) under a technological constraint, the extent to which $u'(c)$ decreases as c increases acts as a redistributive device. That is, the slope of the function $u'(c)$—hence the curvature of $u(c)$—regulates a shift of consumption from well-provided generations to poorer ones, much as a progressive income tax redistributes income among contemporaries. If we again want to exploit the simplicity of (10), we must express also the progressiveness of the redistributing effect of $u'(c)$ in a form unaffected by linear scale

---

* In the present state of our knowledge. For axiomatic discussions of the form of (10) and of some of its alternatives see Koopmans [1960, 1966], Koopmans, Diamond and Williamson [1964], and Diamond [1965].

changes. The expression $1 - u'(c + \gamma)/u'(c)$ for given γ would do, but still depends on an arbitrary γ. This can be avoided by taking

$$\lim_{\gamma \to 0} \frac{1 - u'(c + \gamma)/u'(c)}{\gamma} = -\frac{u''(c)}{u'(c)} = -\frac{d}{dc} \log u'(c) = \eta(c) \quad (13)$$

say, which depends only on c, as it should. The measure $\eta(c)$ will be used later.

As regards the *time horizon* T in (10), for social planning, an infinite horizon $T = \infty$ expresses the fact that no end to society is ever planned. Under present assumptions this creates no complications as long as positive pure time preference is present ($\rho > 0$). But if $\rho = 0$ there is no inherent reason why the utility integral (10) should converge for all paths of interest. Ramsey saved his ethical principle (for a constant population) by the ingenious though somewhat artificial mathematical device of a bliss level $\hat{c}$ of consumption: to exceed that level was by his assumptions either not desired, or not sustainable by the given technology. Instead of maximizing (10), Ramsey then minimized the integral,

$$\int_0^\infty (u(\hat{c}) - u(c_t))\, dt, \quad (14)$$

of the excess of bliss utility over attained utility. This integral converges for the optimal path $\hat{c}_t$ (which satisfies $\lim_{t \to \infty} \hat{c}_t = \hat{c}$) and for all alternative feasible paths worth comparing to it.

Our purpose here is better served if instead of Ramsey's device we employ its modern variant proposed by von Weizsäcker [1965] and named the *overtaking criterion* by Gale [1967]. This criterion achieves the essential comparisons of consumption paths over an infinite future while using only integrals of type (10) for finite values of T. A path c_t is declared better than an alternative path c_t^* if there exists a time T^* such that

$$\int_0^T e^{-\rho t} u(c_t)\, dt > \int_0^T e^{-\rho t} u(c_t^*)\, dt \qquad \text{for all} \quad T \geq T^*. \quad (15)$$

From time T^* onward, the utility integral (10) for path c_t has overtaken that of path c_t^*. The fact that for $\rho = 0$ not every pair of contending paths is comparable under this criterion will turn out to be innocuous.

When the discount rate ρ is positive, use of the overtaking criterion is equivalent to the maximization of (10).

Neither Ramsey nor Harrod indicated in the references cited how the prescription against discounting is to be interpreted if population growth is anticipated. The most highly principled interpretation would seem to require applying the overtaking criterion to

$$\int_0^T L_t u(c_t)\, dt. \tag{16}$$

Here c_t is again per-worker consumption, $u(c_t)$ the utility thereof—or, more precisely, the utility level of each individual, were consumption to be equally distributed among all contemporaries, and were the same utility function applied to all of them. The product $L_t u(c_t)$ then represents the sum of individual utility flows at time t, which (16) integrates over time. Inserting a discount factor $e^{-\rho^* t}$ in (16) would give the criterion

$$\int_0^T e^{-\rho^* t} L_t u(c_t)\, dt, \tag{17}$$

to be called the sum of discounted individual utilities.

In (16) and (17) generations are weighted (before discounting) according to their numbers. An alternative to (17) is to give equal weight to per-worker utilities of different generations, regardless of their size,

$$\int_0^T e^{-\rho t} u(c_t)\, dt. \tag{18}$$

For the same discount rate, $\rho = \rho^*$, these two criteria are obviously quite different. In fact, if the labor force grows by a constant rate λ, as in (1), then the two criteria are mathematically identical if and only if

$$\rho = \rho^* - \lambda. \tag{19}$$

In that case, the criteria are distinct in their interpretation but not in the effects of their implementation.

For definiteness' sake, most of the discussion following is couched in terms of the criterion (18), interpreted literally as the sum of discounted per-worker utilities. However, we shall occasionally use (19) to reinterpret the same findings as applications of (17), the sum of discounted individual utilities. In particular, the existence of this alternative interpretation will lead us to take an interest also in negative values of ρ in (18), which would go against intuition in the literal interpretation.

4. Propositions Concerning Growth Paths Maximizing the Sum of *Un*discounted Per Worker Utilities

Analysis of diagrams with common coordinate axes placed side by side can carry us a long way toward understanding theorems proved elsewhere* about growth paths "optimal" under the various criteria.

In this section, we assume a positive rate of population growth, unless the contrary is specified. In Figures 3 to 8 we consider the objective (18) of a sum of undiscounted per-worker utilities.

The analysis in Section 2 has shown the importance of the function

$$g(k) = f(k) - \lambda k, \tag{20}$$

the excess of the curve $f(k)$ in Figure 1 over the sloping straight line λk. It represents that part of per-worker output available for distribution between per-worker consumption c and net increase $\dot{k}$ in the per-worker capital stock,

$$g(k_t) = c_t + \dot{k}_t. \tag{21}$$

In particular, if during any period k_t is constant, $k_t = k$, then $c = g(k)$ is the constant per-worker consumption resulting therefrom.

Figure 3*a* shows this function with the independent variable k set off on the vertical axis, the values of the function $g(k)$ on the horizontal axis, increasing toward the left. The curve $g(k)$ has a vertical tangent at the point $k = \hat{k}$ corresponding to the golden rule path. Since $g''(k) = f''(k) < 0$ for all k, the curve slopes toward the left for $k < \hat{k}$, toward the right for $k > \hat{k}$, everywhere bending to the right as k increases.

In Figure 3*b*, various alternative paths of capital-per-worker k_t are drawn, with t on the horizontal axis, k_t on the vertical. All paths start, at time $t = 0$, with the given initial per-worker capital stock k_0. We begin by comparing the paths labeled (1) and (2). On path

* The exposition most closely follows Koopmans [1965], where the *i*'s are dotted and the *t*'s are crossed.

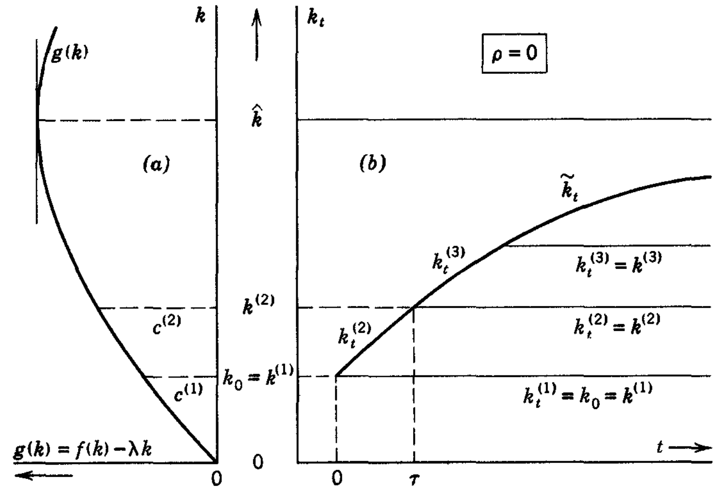

Figure 3 Successively better paths $k_t^{(i)}$, $i = 1, 2, \ldots$.

(1) the per-worker capital $k_t^{(1)} = k_0$ maintains the initial level forever. On path (2), $k_t^{(2)}$ increases over an initial period $0 \leq t < \tau$ to a level $k^{(2)}$ held constant thereafter, $k_t^{(2)} = k^{(2)}$ for all $t \geq \tau$.

Because of the way in which the two diagrams are aligned, the constant per-worker consumption flows $c^{(1)}$, $c^{(2)}$ associated with the level segments of the two paths are read off from the curve $g(k)$ at the levels $k_0 = k^{(1)}$, say, and $k^{(2)}$, respectively. If the initial per-worker capital is below the golden rule level, $k_0 < \hat{k}$, and as long as $k^{(2)} < \hat{k}$ also, we must have $c^{(1)} < c^{(2)}$. This follows from the shape of $g(k)$ already discussed. Since $u(c)$ is increasing with c, we must then have a corresponding relation

$$u^{(1)} = u(c^{(1)}) < u(c^{(2)}) = u^{(2)}$$

for the utility flows on the level portions of the two paths.

On the other hand, over the initial time interval $[0, \tau]$ the investment on path (2) exceeds that on path (1). It is therefore to be expected that this entails a sacrifice of consumption, $c_t^{(2)} < c_t^{(1)}$ for $0 \leq t < \tau$, which is reflected also in the corresponding utility integrals. The various portions of the utility integrals are compared in Table 1. Although we expect $x > 0$, neither the value of x nor that of $y = [u^{(2)} - u^{(1)}]\tau + x$ matters for the outcome of the comparison.

Table 1

	$\int_{\underline{T}}^{\bar{T}} u(c_t)\, dt$ if $[\underline{T}, \bar{T}]$ is		
	$[0, \tau]$	$[\tau, T]$	$[0, T]$
Path (1)	$\tau u^{(1)}$	$(T - \tau)u^{(1)}$	$Tu^{(1)}$
Path (2)	$\tau u^{(1)} - x$, say	$(T - \tau)u^{(2)}$	$Tu^{(2)} - y$, say

By the overtaking criterion we must determine whether, for large enough T,

$$\int_0^T (u(c_t^{(2)}) - u(c_t^{(1)}))\, dt = (u^{(2)} - u^{(1)})T - y$$

is positive. Since $u^{(2)} - u^{(1)} > 0$, this is the case for all $T \geq T^*$ if

$$T^* = \text{the larger of the numbers 1 and } \frac{y}{u^{(2)} - u^{(1)}} + 1,$$

surely a finite positive number. Hence path (2) is better than path (1).

Note that this reasoning is independent of the length of the time interval $[0, \tau]$ and of the level $k^{(2)}$ at which path (2) becomes constant, as long as $k^{(1)} < k^{(2)} < \hat{k}$. Therefore path (3) is again better than path (2), and so on. Thus, given any path such as $\bar{k}_t$ in Figure 3, which rises from k_0 and either approaches the golden rule level $\hat{k}$ as an asymptote, or attains, and remains at, that level from a certain point in time on, we find that any path initially coinciding with $\bar{k}_t$ and then branching off to remain constant at some level below $\hat{k}$ is overtaken by any other such path that branches off later, at a higher level below $\hat{k}$.

If, on the other hand, $k_0 > \hat{k}$, a similar result is obtained in which the word "rises" is replaced by "falls," "below" by "above," and "higher" by "lower."

These comparisons are made within a highly restricted class of paths. Could a path that fluctuates, finitely or infinitely often, be better than any path that moves in one direction or stays put?

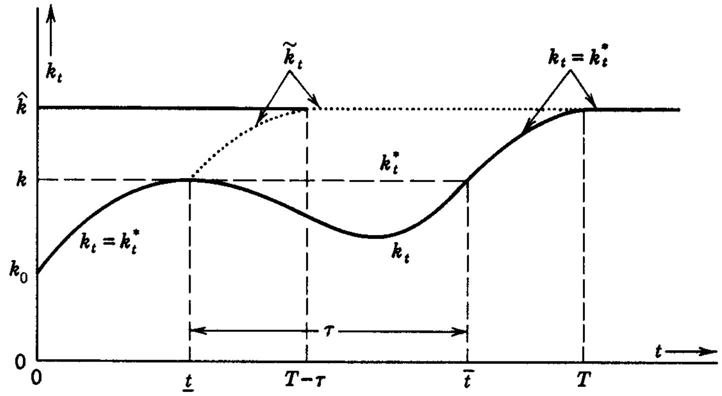

Figure 4 Nonoptimality of bulges and of level segments with $k \neq \hat{k}$.

Figure 4 shows that a path k_t that has at least one fluctuation, let us say extending *below* the golden rule level $\hat{k}$, must contain a *bulge*. This is defined as a time interval $[\underline{t}, \bar{t}]$ in which k_t attains the *same* below-golden-rule level k at its beginning and its end, $k_{\underline{t}} = k = k_{\bar{t}} < \hat{k}$, and *lower* levels for $\underline{t} < t < \bar{t}$. Compare k_t with a path k_t^* remaining constant at $k_t^* = k_{\underline{t}}$ for $\underline{t} \leq t \leq \bar{t}$, and coinciding with k_t at all other times. Then, over the interval $[\underline{t}, \bar{t}]$, $g(k_t)$ averages less than $g(k_t^*)$. Over the same interval, the net increase of k_t, as well as that of k_t^*, equals zero, hence averages zero. Therefore, by (21), c_t averages less than c_t^*. But since $u(c)$ is concave and c_t fluctuates, $u(c_t)$ averages at less than u(average of c_t), whereas $u(c_t^*)$ averages to u(average of c_t^*) because c_t^* is constant. Therefore k_t^* has overtaken k_t from $t = \bar{t}$ on. Similar reasoning precludes mirror image bulges above $\hat{k}$.

Not even a flat segment at a level k different from $\hat{k}$ can be part of an optimal path. Figure 4 further compares the above path k_t^*, which is now assumed to attain the golden rule level $\hat{k}$ from $t = T$ on, with another path $\tilde{k}_t$ defined by

$$\tilde{k}_t = k_t, \qquad 0 \leq t \leq \underline{t},$$

$$\tilde{k}_t = k_{t+\tau}, \qquad \underline{t} \leq t, \qquad \tau = \bar{t} - \underline{t}$$

which anticipates the post-$\bar{t}$ future course of k_t immediately following $\underline{t}$, thus omitting the flat segment, and attaining $\hat{k}$ from time $T - \tau$ on.

Table 2

	$\int_{\underline{T}}^{\bar{T}} u(c_t)\,dt$ if $[\underline{T}, \bar{T}]$ is	$\int_{\underline{T}}^{\bar{T}} u(c_t)\,dt$ if $[\underline{T}, \bar{T}]$ is	$\int_{\underline{T}}^{\bar{T}} u(c_t)\,dt$ if $[\underline{T}, \bar{T}]$ is	$\int_{\underline{T}}^{\bar{T}} u(c_t)\,dt$ if $[\underline{T}, \bar{T}]$ is	$\int_{\underline{T}}^{\bar{T}} u(c_t)\,dt$ if $[\underline{T}, \bar{T}]$ is
	$[\underline{t}, \bar{t}]$	$[\bar{t}, T]$	$[\underline{t}, T - \tau]$	$[T - \tau, T]$	$[\underline{t}, T]$
Path k_t^*	$\tau u(g(k))$	x, say			$\tau u(g(k)) + x$
Path $\tilde{k}_t$			x, again	$\tau u(g(\hat{k}))$	$\tau u(g(\hat{k})) + x$

The comparison is made in Table 2, omitting those parts of the future for which the two paths coincide. From time T on, path $\tilde{k}_t$ has overtaken path k_t^* by $\tau(u(g(\hat{k})) - u(g(k)))$, a positive amount whenever $k \neq \hat{k}$. As will be seen in another case, this reasoning can be refined for a path k_t that approaches the level $\hat{k}$ asymptotically instead of attaining it at some finite time.

We now know that, if an optimal path $\hat{k}_t$ exists, it must approach $\hat{k}$ monotonically in a finite or infinite time. We can determine the required shape of $\hat{k}_t$ if we can find out how the slope $\dot{\hat{k}}_t$ of the path $\hat{k}_t$ depends on the level attained at time t. There is no loss of generality in looking at this problem just for time $t = 0$, for various alternative values of k_0.

This time the question raised cannot be answered without bringing the shape of the utility function $u(c)$ into the diagram. A beautifully simple reasoning, suggested by Keynes to Ramsey [1928] for the case of a constant population, can readily be adapted to the present case of population growth. It is one of these intuitive heuristic arguments that convey the simple answer in a flash to a reader willing to be persuaded which quantities of "first order of smallness" need to be carried along and which quantities of "higher order of smallness" can be ignored.

We first cite Ramsey's rendering of Keynes' argument, changing the notations to correspond to those used here, and omitting Ramsey's reference to the disutility of labor.

"Suppose that in a year we ought to spend £c and save £s. Then the advantage to be gained from an extra £1 spent is $u'(c)$, the marginal utility of money, and this must be related to the sacrifice imposed by saving £1 less.

"Saving £1 less in the year will mean that we shall only save £s in $1 + 1/s$ years, not, as before, in one year. Consequently, we shall be in $1 + 1/s$ year's time exactly where we should have been in one year's time, and the whole course of our approach to bliss will be postponed by $1/s$ of a year, so that we shall enjoy $1/s$ of a year less bliss and $1/s$ of a year more at our present rate. The sacrifice is, therefore,

$$(1/s)(u(\hat{c}) - u(c)).$$

Equating this to $u'(c)$ we get

$$\dot{k} = f(k) - c = \frac{u(\hat{c}) - u(c)}{u'(c)},$$

if we replace s by $\dot{k}$, its limiting value."

The following paragraphs apply this reasoning in somewhat greater detail to the case of population growth, where the asymptote of $\hat{k}_t$ is the golden rule level $\hat{k}$ rather than a bliss level.

Assume that a smooth optimal path $\hat{k}_t$ as shown in Figure 5*a* exists. For the moment the datum is the initial per-worker capital k_0, the unknown its initial rate of increase $\dot{k}_0$, the slope of $\hat{k}_t$ at $t = 0$. Choose a time unit small enough that, on the interval [0, 1], $\hat{k}_t$ can be treated as a straight line segment, hence the variation of $\dot{k}_t$ can be ignored. The time interval should also be small enough so that the variation of $g(k(t))$ for $0 \leq t \leq 1$ can be ignored. Then, at $t = 1$, per-worker capital "equals" $k_0 + \dot{k}$, whereas per-worker consumption up to that time runs to

$$\hat{c}_0 \approx g(k_0) - \dot{k}_0,$$

the consumption $c_0 = g(k_0)$ that would have occurred in one unit of time had k_t remained constant, less the actual increment $\dot{k}_0$ to k_t in that time. The numbers c_0 and $\hat{c}_0$ are transferred, with the help of a mirror suitably positioned in Figure 5*b* at a 45° slope, to the c-axis in Figure 5*c*.

In tracing utility implications of alternative paths in Figure 5*c* we use our option to change the utility scale linearly by adopting the scale

$$v(c) = u(c) - u(\hat{c}) = u(c) - \hat{u},$$

in which the golden rule consumption level $\hat{c}$ produces a zero flow of v-utility. The v-utility accumulated in the first unit of time along the path $\hat{k}_t$ is then $\hat{v}_0 = v(\hat{c}_0)$ as shown in Figure 5*c*. To make clear that this is the product of a rate $\hat{v}_0$ with the length 1 of a time interval,

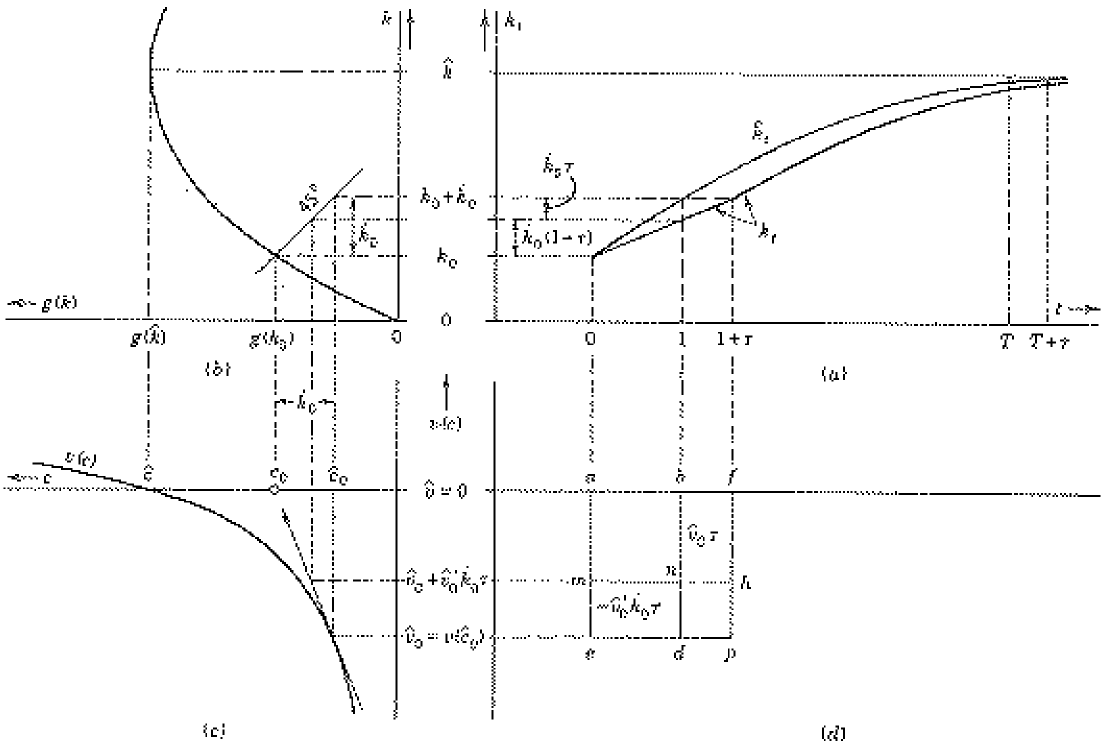

Figure 5 Determination of $\dot{k}_0$ and $\hat{c}_0$.

we represent it by the area of the rectangle *abde* in Figure 5*d*,

$$A(abde) = \hat{v}_0 \cdot 1.$$

(Because it is below the horizontal axis, this area is to be counted as a negative number.)

Next we choose a τ which in turn is absolutely small compared with 1, and compare $\hat{k}_t$ with a path k_t which attains the level $\hat{k}_1 = k_0 + \dot{k}_0$ at the slightly different time $1 + \tau$ (slightly later if $\tau > 0$), while following a straight line path up to that time. Thereafter, k_t imitates $\hat{k}_t$ with a delay τ,

$$k_t = \hat{k}_{t-\tau}, \qquad t \geq 1 + \tau.$$

Then, on the interval $[0, 1 + \tau]$, the rate of increase in k_t is

$$\frac{\dot{k}_0}{1 + \tau} \approx \dot{k}_0(1 - \tau) = \dot{k}_0 - \dot{k}_0\tau,$$

consumption flows at the rate $\hat{c}_0 + \dot{k}_0\tau$, v-utility at the rate

$$v(\hat{c}_0 + \dot{k}_0\tau) \approx \hat{v}_0 + \hat{v}_0'\dot{k}_0\tau,$$

taking the tangent, with slope $\hat{v}_0' = v'(\hat{c}_0)$, as if it were the curve; v-utility accruing over the period $[0, 1 + \tau]$ is therefore

$$A(afhm) = (\hat{v}_0 + \hat{v}_0'\dot{k}_0\tau)(1 + \tau).$$

Finally, we choose T so large that $\hat{k}_T$ has become equal to, or at least "equal" to, the golden rule level,

$$\hat{k}_T = k_{T+\tau} \approx \hat{k},$$

so that any remaining difference between $\hat{k}_t$ and k_t for $t \geq T$ can be ignored. Table 3 shows the comparison of v-utility accruals. By

Table 3

	$\int_{\underline{T}}^{\overline{T}} v(c_t)\,dt$ if $[\underline{T}, \overline{T}]$ is					
	$[0, 1]$	$[1, T]$	$[T, T+\tau]$	$[0, 1+\tau]$	$[1+\tau, T+\tau]$	$[0, T+\tau]$
Path $\hat{k}_t$	$A(abde) = \hat{v}_0 \cdot 1$	x, say	$\hat{v}\tau = 0$			$\hat{v}_0 \cdot 1 + x$
Path k_t				$A(afhm) = (\hat{v}_0 + \hat{v}_0'\dot{k}_0\tau) \cdot (1 + \tau)$	x	$\hat{v}_0 + \hat{v}_0'\dot{k}_0\tau + \hat{v}_0\tau + x$

time $T + \tau$, path k_t is "ahead" of path $\hat{k}_t$ by

$$A(afhm) - A(abde) \approx A(bfpd) - A(mnde) = (\hat{v}_0 + \hat{v}_0'\dot{k}_0)\tau \quad (22)$$

(throwing in $A(nhpd)$, proportional to τ^2, for good measure). Although τ must be small in absolute value, it can be either positive or negative. Hence, if the coefficient of τ in (22) were to be different from zero, an absolutely small enough τ of the same sign would make k_t slightly better than $\hat{k}_t$. The vanishing of the coefficient of τ in (22) is therefore a necessary condition for the optimality of $\dot{k}_0$ as the slope of $\hat{k}_t$ at $t = 0$,

$$\hat{v}_0 + \hat{v}_0'\dot{k}_0 = 0 \quad (23)$$

Geometrically, this says precisely that the tangent to the curve $v(c)$ in Figure 5c at the point $\hat{c}_0$ must pass through the point $(c, v) = (c_0, 0)$. (To see this, let τ approach 1 in 5b and 5c.) Reversing the reasoning, the construction of the optimal initial consumption rate $\hat{c}_0$ proceeds from the given k_0 via the curve $g(k)$ in 5b to the point marked c_0 in 5c, from which a tangent to the curve $v(c)$ is drawn, with $\hat{c}_0$ as the c-coordinate of the tangency point. Then $\dot{k}_0 = c_0 - \hat{c}_0$.

As stated previously, the same construction applies to determining the optimal slope $\dot{k}_t$ from any given value $\hat{k}_t$ reached by the optimal capital path at some given other time t. Reverting to the original utility scale, we have therefore found

$$\dot{k}_t = \frac{u(\hat{c}) - u(\hat{c}_t)}{u'(\hat{c}_t)} \quad (24)$$

to be the differential equation connecting any jointly optimal consumption and capital paths. For the determination of both paths from a given k_0, (24) has to be combined* with the identity (21) incorporated in the construction of Figure 5.

Note that, in spite of changes in the interpretations of the variables, (24) has the same form as the condition derived by the Keynes-Ramsey reasoning.

Figure 6a suggests how the values of $\dot{k}_0$ vary with alternative (unlabeled) initial values of k_0. It also illustrates how the slope $\dot{k}_{t'}$ of the optimal path $\hat{k}_t$ at any time $t = t'$ can be read off from the

* Since elimination of $\dot{k}_t$ from (21) and (24) produces a relation between $\hat{c}_t$ and k_t directly, the optimal paths are determinable from one differential equation of the first order.

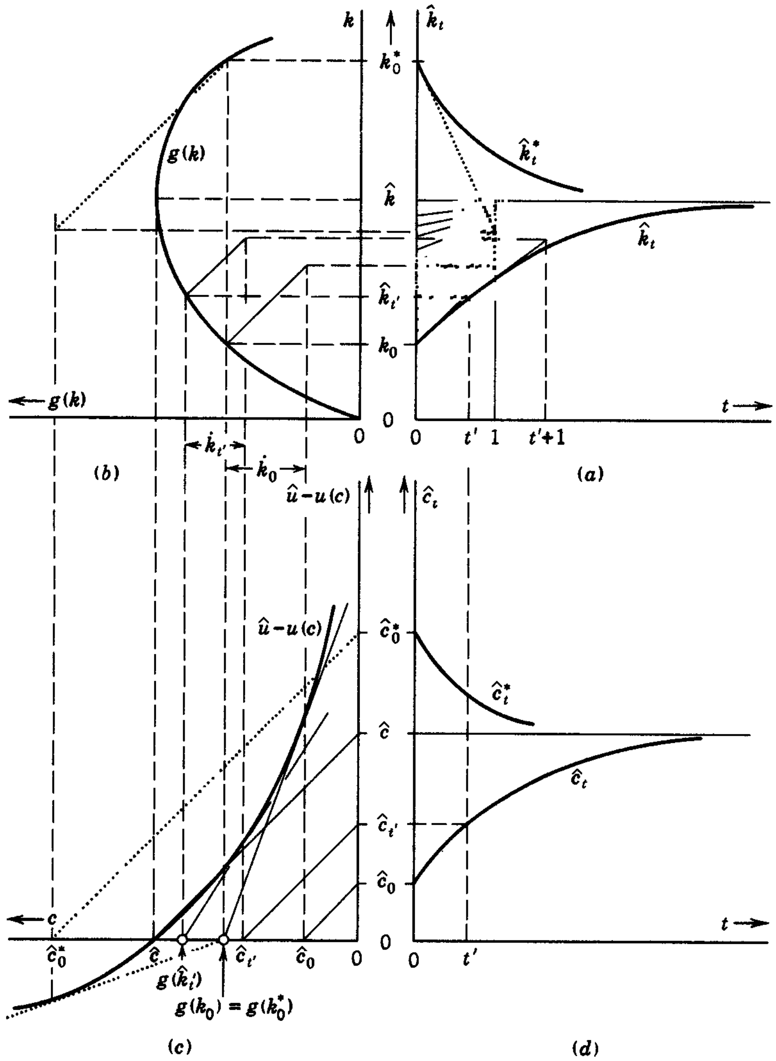

Figure 6 Construction of optimal paths $\hat{k}_t$, $\hat{c}_t$.

diagram. Furthermore, using the negative $\hat{u} - u(c) = -v(c)$ of the v-utility function used in Figure 5, it indicates how the optimal consumption rates $\hat{c}_0, \hat{c}_{t'}, \ldots,$ determined from tangency points can be transferred from Figure 6*c* to 6*d* (using lines of 45° slope) to construct the entire optimal consumption path $\hat{c}_t$ associated with $\hat{k}_t$. It follows from the construction that $\hat{c}_t$ rises monotonically to approach the golden-rule per-worker consumption level as an asymptote.

Finally, Figure 6 shows how an initial per-worker capital k_0^* just enough *larger* than the golden rule value $\hat{k}$ to make $g(k_0^*) = g(k_0)$ leads to a construction of $\hat{k}_0^*$ and $\hat{c}_0^*$ based on the *other* tangent to the curve in 6*c*, drawn out of the same point $c_0 = g(k_0)$, using entirely similar reasoning.

Finding a unique pair of paths $\hat{k}_t$, $\hat{c}_t$ jointly meeting necessary conditions for optimality does not prove their optimality. It has been shown elsewhere* that the pair of paths meeting these conditions is indeed optimal, and that the golden-rule levels $\hat{k}$, $\hat{c}$, are approached only asymptotically.

If one lets the growth rate of the labor force approach zero, then under present assumptions the golden-rule capital stock $\hat{k}$ approaches infinity; so the present solution evaporates. However, if for $\lambda = 0$ we adopt Ramsey's assumption that the per-worker production function $f(k)$ [now $= g(k)$] reaches a maximum for a finite per-worker capital stock $\hat{k}$ (capital saturation), then our solution reverts to the Keynes-Ramsey formula: *Along an optimal path the rate of saving* (= investment) *equals the excess of the maximum sustainable utility level over the utility of the present optimal rate of consumption, divided by the marginal utility of consumption at the latter rate.* We have phrased this rule in such a way that it can be applied to positive rates of labor force growth as well, by the mere insertion of the adjectival "per-worker" in suitable places (and by interpreting "per-worker investment" as $\dot{k}_t$, the rate of increase in per-worker capital).

5. Comparative Dynamics

The diagrammatic procedures developed in Section 4 can be used to study how the pair $\hat{k}_t$, $\hat{c}_t$ of optimal paths changes if one varies

* Koopmans [1965], Proposition (C), and Inagaki [July, 1966].

the production function $f(k)$, the utility function $u(c)$, or the discount rate ρ, in some given manner.

Effect of the Marginal Productivity of Capital

One would expect that, in comparing two production functions f, f^* with the same per-worker output $f(k_0) = f^*(k_0)$ at the initial per-worker capital k_0, but different marginal productivities

$$f'(k_0) > f'^*(k_0), \tag{25}$$

the smaller marginal productivity would, by diminishing the future increments in consumption attainable through an extra unit of present investment, lead to a larger consumption in the present. This is confirmed by Figure 7, where $f(k)$ and $f^*(k)$ have been chosen to lead to the *same* golden-rule per-worker capital $\hat{k} = \hat{k}^*$, but different golden-rule per worker consumption rates,

$$\hat{c} = g(\hat{k}) > g^*(\hat{k}) = \hat{c}^*. \tag{26}$$

Rather than drawing two different parallel curves $\hat{u} - u(c)$ and $\hat{u}^* - u(c)$ in Figure 7c, we draw one curve and refer it to two different vertical scales, identified by the origins O, O^*, respectively. Then the point $(g(k_0), 0)^*$ referred to O^* is vertically above the point $(g(k_0), 0)$ referred to O. In view of the curvature of the graph of $\hat{u} - u(c)$, the point of tangency determining $\hat{c}_0^*$ then is necessarily to the left of that determining $\hat{c}_0$, so we have

$$\hat{c}_0^* > \hat{c}_0, \qquad k_0^* < k_0,$$

as anticipated. The latter inequality is read off in Figure 7a, ignoring the discrepancies at $t = 1$ between the curves $\hat{k}_t$, $\hat{k}_t^*$, and their respective tangents at $t = 0$.

Further analysis shows that $\hat{c}_t^*$ falls below $\hat{c}_t$ from some positive t' on, for two reasons. In the first place, since $g^*(k)$ represents a less productive technology than $g(k)$ for higher capital intensities $k > k_0$, $\hat{c}_t$ has by (26) a higher asymptote than c_t^*. In addition to this, if in the technology $g(k)$ a path $\tilde{c}_t$ started out with the initial consumption rate $\tilde{c}_0 = \hat{c}_0^* > \hat{c}_0$, it would on feasibility grounds alone have to pay for this higher immediate consumption by lower rates of consumption, $\tilde{c}_t < \hat{c}_t$, at some later time.

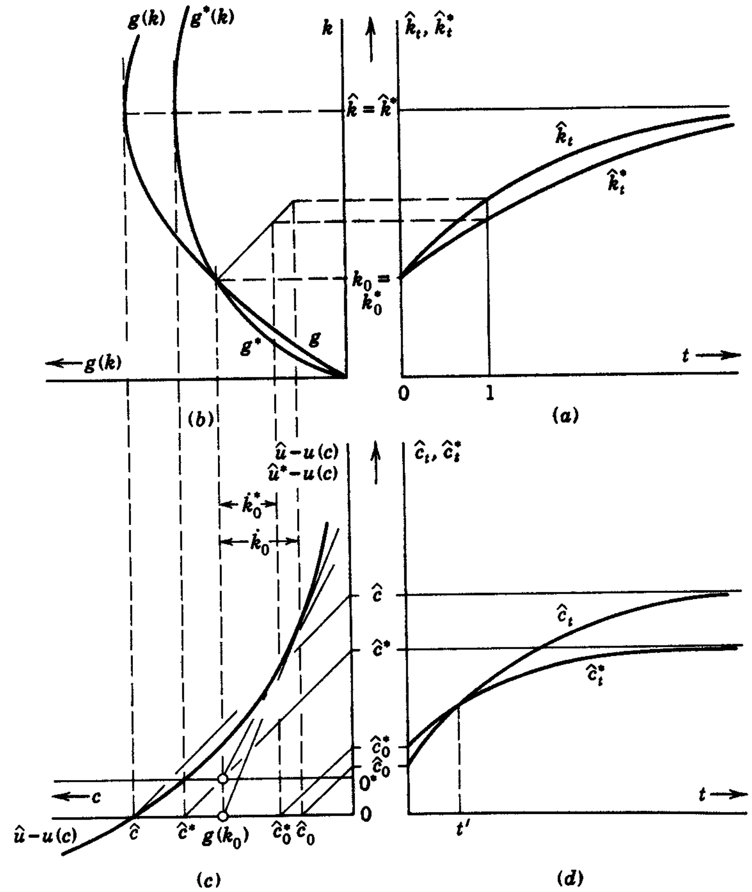

Figure 7 Effect of $g'(k_0)$ on optimal paths.

The two optimal capital paths, $\hat{k}_t$ and $\hat{k}^*_t$, have the same asymptote $\hat{k} = \hat{k}^*$ by our assumption, with $\hat{k}^*_t$ trailing behind $\hat{k}_t$ at least initially.

Effect of the "Curvature" of the Utility Function

It was observed in Section 3 that the "curvature" $\eta(c)$ of the utility function affects the distribution of consumption between periods of markedly different rates of consumption. In Figure 8*c*, the curve

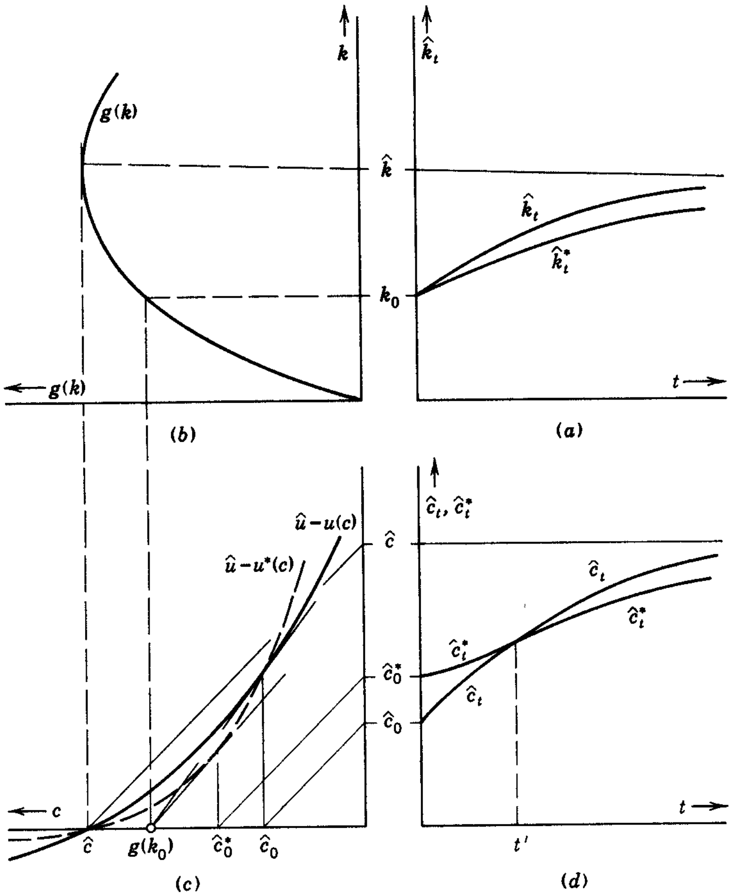

Figure 8 Effect of the "curvature" of the utility function.

$\hat{u} - u(c)$ is contrasted with a curve $\hat{u} - u^*(c)$ which, whatever its original scale, has been so (linearly) rescaled that the two curves intersect precisely at the golden-rule consumption rate $\hat{c}$ and again at the consumption rate $\hat{c}_0$ optimal if $u(c)$ defines the criterion of optimality,

$$\hat{u} = u(\hat{c}) = u^*(\hat{c}), \qquad u(\hat{c}_0) = u^*(\hat{c}_0).$$

It then is immediately apparent that the more highly curved $u^*(c)$

leads to the higher consumption rate $\hat{c}_t^*$ at the time $t = 0$ (and for some time thereafter), when per-worker consumption is, in both paths, relatively low. Since this also causes $\hat{k}_t$ to rise above $\hat{k}_t^*$, the consumption paths $\hat{c}_t$, $\hat{c}_t^*$ are bound to cross at some later time t'.

Effect of Discounting

To discuss the effect of a *positive* discount rate ρ, we revert to the type of analysis of Figure 4 in which only the monotonicity, not the shape of $u(c)$, is used, and only the monotonicity and the asymptote, not otherwise the shape of $\hat{k}_t$ are determined.

Figure 9 compares paths similar to those of Figure 3, but differs only in that it applies the utility functional (18) with a positive value of ρ. Using the same notations as before, we now specify that *both* τ *and the slope* $d\tilde{k}_t/dt$ *of the* tentative rising *capital path* shall be small. Small τ allows us to ignore discounting on $[0, \tau]$. If the slope of $\tilde{k}_t$ is also small, differences in utility flows between all segments of paths to be compared are small enough for us to replace the utility curve $u(c)$ by its tangent at the point $c^{(1)} = g(k_0)$. The criterion can then be simplified to the integral over discounted rates of consumption instead of the associated utility flows, as shown in Table 4. The difference between the first column entries for paths (1) and (2) arises, of course, from the additional investment made on path (2).

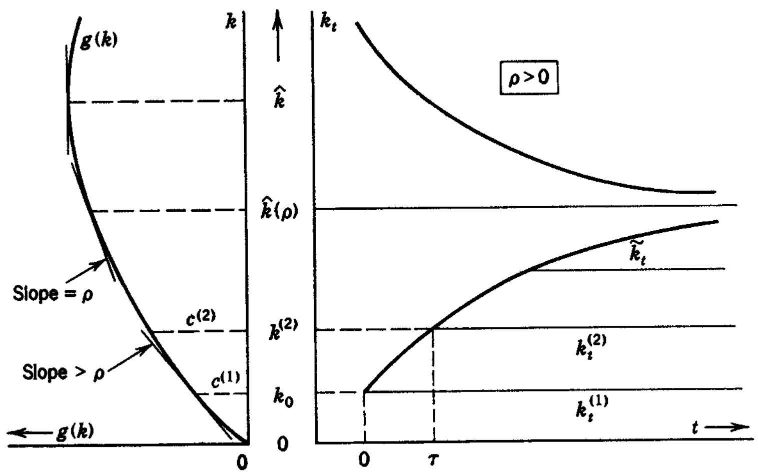

Figure 9 Effect of discounting per-worker utilities ($\rho > 0$).

Table 4

	$\int_{\underline{T}}^{\bar{T}} e^{-\rho t} c_t \, dt$ if $[\underline{T}, \bar{T}]$ is	
	$[0, \tau]$	$[\tau, \infty]$
Path (1)	$\tau c^{(1)}$	$(e^{-\rho\tau}/\rho)c^{(1)}$
Path (2)	$\tau c^{(1)} - [k^{(2)} - k^{(1)}]$	$(e^{-\rho\tau}/\rho)c^{(2)}$
Excess, (2) over (1)	$-[k^{(2)} - k^{(1)}]$	$[c^{(2)} - c^{(1)}]/\rho$

The second column entries are

$$c^{(i)} \int_{\tau}^{\infty} e^{-\rho t} \, dt = c^{(i)} \left(\frac{e^{-\rho\tau}}{\rho} \right), \qquad i = 1, 2.$$

Since $c^{(2)} - c^{(1)}$ is itself of the order of τ, the difference between $e^{-\rho t}$ and 1 can be ignored in the last entry of the table.

Path (2) is better than path (1) if the sum of the entries in the last row is positive, that is, if

$$\frac{c^{(2)} - c^{(1)}}{k^{(2)} - k^{(1)}} > \rho. \tag{27}$$

This says, understandably enough, that the ratio of the additional perpetual *per capita* consumption flow $c^{(2)} - c^{(1)}$ to the initial *per capita* consumption sacrifice $k^{(2)} - k^{(1)}$ that made it possible must exceed the discount rate applicable to *per capita* utility. Figure 9 shows that this will be the case as long as both $k^{(1)}$ and $k^{(2)}$ stay below that value $\hat{k}(\rho)$ for which

$$g'[\hat{k}(\rho)] = \rho, \qquad \text{so } f'[\hat{k}(\rho)] = \rho + \lambda.$$

We conclude that, if $\bar{k}_t$ is a path rising *sufficiently slowly* from k_0 to an asymptotic level $\hat{k}(\rho)$, then among the paths branching off from $\bar{k}_t$ to remain constant from some time t' on, the path branching off later is always better. (In this case, the pertinent integrals converge on the interval $[0, \infty)$, and the overtaking criterion and the maximization of the utility functional (18) on $[0, \infty)$ give the same answer.)

For the pair of paths $(\tilde{k}_t, \tilde{c}_t)$ to be optimal, it must now satisfy a system of two differential equations of the first order examined elsewhere*.

As explained previously, the optimal per-worker capital and consumption paths found by maximizing the sum (18) of per-worker utilities discounted at a rate $\rho \geq 0$ can also serve as optimal paths with reference to the sum of individual utilities discounted at a rate $\rho^* = \rho + \lambda \geq \lambda$.

6. The Splurge that Gains from Postponement

The examples of Section 5 have indicated how "optimal" intertemporal distribution depends on specific traits of the production function and of the utility functional. In particular, we have seen that posterity is favored, *ceteris paribus*, by a high marginal productivity of capital, by a low discount rate, and—if initial capital falls short of the golden-rule level—by low "curvature" of the utility function. The point to be made in this section, again by an example, is that slanting the data of technology or of policy too much in favor of posterity can be self-defeating. We shall show this by considering a negative discount rate, $\rho < 0$, as applied to per-worker utilities. As explained, this can be more naturally interpreted as the case in which a discount rate

$$\rho^* < \lambda,$$

smaller than the rate of labor force growth is applied to individual utilities (17) before their summation.

In Figure 10, we consider a long but finite horizon T, and specify (just to choose something) that the terminal per-worker capital shall be at the golden-rule level,

$$k_T = \hat{k}.$$

We shall argue that the path $\hat{k}_t$ "optimal" under that additional constraint will bulge out as shown and will, if T is large enough, spend most of the period close to that level $k(\rho)$ where the tangent to the function $g(k)$ has the (now negative) slope ρ.

* Koopmans [1965], Propositions (I), (J), and Section A. 7.

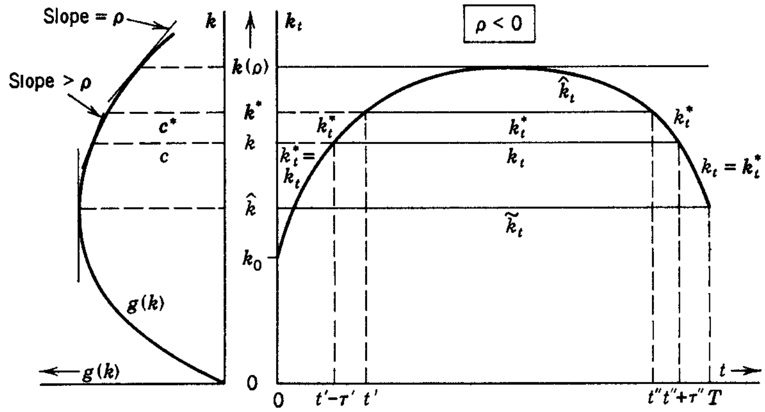

Figure 10 Effect of discounting at a negative rate ρ.

Compare the paths k_t, k_t^*, that short-cut the bulging curve $\hat{k}_t$ by level stretches at levels $k = \hat{k}_{t'-\tau'} = \hat{k}_{t''+\tau''}$ and $k^* = \hat{k}_{t'} = \hat{k}_{t''}$, respectively. Take $k^* > k$, but take the difference $k^* - k$ so small that the variation in the discount factor can be ignored within $[t' - \tau', t']$, and again within $[t'', t'' + \tau'']$ and that $u(c)$ can again, on $[c, c^*]$, be replaced by its tangent. Writing $\sigma = -\rho > 0$ for the negative of the discount rate, Table 5 compares the discounted

Table 5

	$\int_{\underline{T}}^{\bar{T}} e^{\sigma t} c_t \, dt$ if $[\underline{T}, \bar{T}]$ is		
	$[t' - \tau', t']$	$[t', t'']$	$[t'', t'' + \tau'']$
Path c_t	x, say	$\frac{c}{\sigma}(e^{\sigma t''} - e^{\sigma t'})$	y
Path c_t^*	$x - (k^* - k)e^{\sigma t'}$	$\frac{c^*}{\sigma}(e^{\sigma t''} - e^{\sigma t'})$	$y + (k^* - k)e^{\sigma t''}$
Excess, c_t^* over c_t	$-(k^* - k)e^{\sigma t'}$	$\frac{c^* - c}{\sigma}(e^{\sigma t''} - e^{\sigma t'})$	$(k^* - k)e^{\sigma t''}$

exist recurs in more general models. In models with exponential technical progress, product-augmenting (Inagaki) or labor-augmenting (Mirrlees), similar critical points have been found that depend on the rate of progress, on the shape of the utility function for large rates of consumption, and, if progress is product-augmenting, on the shape of the production function for large capital-labor ratios.

7. Concluding Remarks

The moral of our story is that ethical principles, in the subject-matter in hand, need mathematical screening to determine whether in given circumstances they are capable of implementation. Only principles that have passed such a test present ethical, or policy, problems.

More specifically, the maximization in a constant technology of the sum of discounted utilities of all future members of an exponentially growing population makes sense only if the discount rate at least equals the rate of population growth. Failing that, the maximization cannot be carried through with an infinite horizon. With an arbitrary choice of a finite horizon, the principal feature of the "optimal" path completely depends on that arbitrary choice.

REFERENCES

Böhm-Bawerk, E. von [1921], *Positive Theory of Capital*, Vol. II of *Capital and Interest*, 1959, Libertarian Press, translated from fourth edition of original in German, 1921.

Cass, D. [1964], "Optimum Growth in an Aggregate Model of Capital Accumulation: Turnpike Theorem," *Cowles Foundation Discussion Paper* 178, November, 1964, to be published in *Econometrica*.

Diamond, P. A. [1965], "The Evaluation of Infinite Utility Streams," *Econometrica*, January, 1965, pp. 170–177.

Fisher, I. [1930], *The Theory of Interest*, reprinted in 1961, Augustus Kelley, New York.

Gale, D. [1967], "On Optimal Development in a Multi-Sector Economy," *Review of Economic Studies*, January, 1967, pp. 1–18.

Harrod, R. F. [1948], *Towards a Dynamic Economics*, Macmillan, 1948.

---

* Inagaki [1966], Koopmans [1967], Mirrlees [1967].

Inagaki, M. [1966], "Utility Maximization over Infinite Time: A General Existence Theorem," Netherlands Economic Institute, Division of Balanced International Growth, Publ. No. 34/66, February, 1966; "Utility Maximization: An Explicit Solution," Discussion Paper, May, 1966, and "Utility Maximization under Constant Technology," July, 1966 (all mimeographed).

Koopmans, T. C. [1960], "Stationary Ordinal Utility and Impatience," *Econometrica*, April, 1960, pp. 287–309.

——— [1965], "On the Concept of Optimal Economic Growth," in *The Econometric Approach to Development Planning*, North-Holland Publ. Co. and Rand-McNally, 1966, a reissue of *Pontificiae Academiae Scientiarum Scripta Varia*, Vol. 28, 1965, pp. 225–300.

——— [1966], "Structure of Preference Over Time," *Cowles Foundation Discussion Paper No. 206*, April, 1966, 42 pp.

——— [1967], "Objectives, Constraints and Outcomes in Optimal Growth Models," *Econometrica*, January, 1967.

———, P. A. Diamond and R. E. Williamson [1964], "Stationary Utility and Time Perspective," *Econometrica*, January-April 1964, pp. 82–100.

Mirrlees, J. [1967], "Optimum Growth when Technology is Changing," *Review of Economic Studies*, January, 1967, pp. 95–124.

Phelps, E. S. [1966], *Golden Rules of Economic Growth*, Norton 1966.

Ramsey, F. P. [1928], "A Mathematical Theory of Saving," *Economic Journal*, December, 1928, pp. 543–559.

Samuelson, P. A. [1965], "A Catenary Turnpike Theorem Involving Consumption and the Golden Rule," *American Economic Review*, June, 1965, pp. 486–496.

Weizsäcker, C. C. von [1965], "Existence of Optimal Programmes of Accumulation for an Infinite Time Horizon," *Review of Economic Studies*, April, 1965, pp. 85–104.