2

Introduction to Dynamic Optimization Theory

Tapan Mitra

1. Introduction

Dynamic optimization models and methods are currently in use in a number of different areas in economics, to address a wide variety of issues. The purpose of this chapter is to provide an introduction to the subject of dynamic optimization theory which should be particularly useful in economic applications. The current section provides a brief history of our subject, to put this survey in proper perspective. It then provides an overview of the topics covered in the various sections of this review.

The contribution by Frank Ramsey (1928) on the optimum rate of savings of a nation is generally regarded as the paper which introduced the use of dynamic optimization methods in addressing economic problems. Ramsey analyzed a continuous-time dynamic optimization model, and developed a modification of the standard calculus of variations method to deal with the problem of existence of an optimum savings rate, when all generations (current and future) over an infinite horizon are to be treated equally in the objective function.

Almost thirty years were to pass before there was a renewed interest in the problem of optimum savings as formulated by Ramsey. The method of optimal control formulated by Pontryagin et.al. (1962), had a significant impact on this line of research. Ramsey's problem was reformulated and studied in depth in terms of the continuous-time one-sector neoclassical model by Cass (1965), Koopmans (1965) and Samuelson (1965), and in the continuous-time two-sector neoclassical model by Srinivasan (1964) and Uzawa (1964).

The 1950s had seen the successful application of the more basic

M. Majumdar et al., *Optimization and Chaos*

methods of real and convex analysis in general equilibrium theory by Arrow-Debreu (1954), McKenzie (1954), Gale(1955) and Nikaido(1956), replacing the earlier calculus treatments of the problem of existence of equilibrium and its Pareto-Optimality. In his seminal work on efficient allocation of resources over time in an infinite-horizon framework, Malinvaud (1953) recognized that problems of intertemporal allocation could be addressed by employing similar methods, thereby bringing the treatment of capital theory closer to that of the theory of general equilibrium. In his justly celebrated survey of the state of economic science, Koopmans (1957) reviewed these developments and advocated the use of these methods more generally in all branches of economic theory. The impact of this line of thinking on the theory of dynamic optimization took some time to materialize. The standard problem of dynamic optimization was formulated both as a discrete-time problem, and in alternative versions of the so-called reduced form model, by Radner (1967a), using dynamic programming methods, and by Gale (1967) and McKenzie (1968), using the methods of duality theory. Gale's paper appeared along with the papers by McFadden (1967) and Radner (1967b) in a symposium of the Review of Economic Studies, which had a substantial impact on the methods subsequently used in dynamic optimization theory.

The papers by Gale (1967) and McKenzie (1968) were concerned, true to the spirit of the Ramsey exercise, with objective functions in which future utilities were not discounted. Infinite-horizon programs were compared by means of some version of the "overtaking criterion" proposed by Atsumi (1965) and Weizsacker (1965). The central problem was seen as the problem of existence of an optimal program, and its dynamic behavior (convergence to the unique stationary optimal stock) was a by-product, obtained en route to solving this central problem.

The 1950s and early sixties had given rise to a literature on finite-horizon pure capital accumulation oriented dynamic optimization exercises, where optimality was defined in terms of only the state of the economy at the end of the horizon. Samuelson (1949) had conjectured that programs, optimal according to this criterion, would stay close (for most of the planning horizon) to the balanced growth path which had the largest growth factor, associated with a von Neumann equilibrium (1945). This was made more explicit in Dorfman, Samuelson and Solow (1958) and came to be known as the "turnpike conjecture". Definitive proofs of this conjecture were obtained in alternative frameworks by Morishima(1961), Radner(1961) and McKenzie(1963a). The central idea contained in Radner's paper was that optimal programs which stayed away (uniformly) from the von Neumann equilibrium would suffer a (uniform) value loss, and McKenzie (1963b, 1963c, 1971) recognized that this was the key concept to generalizations of the turnpike property.

When the interest of the profession shifted from purely capital accumulation oriented models to consumption oriented optimal growth models of the Ramsey-type, this key concept remained. Both Atsumi (1965) and McKenzie (1968) recognized that this idea could be used to advantage in studying asymptotic properties of optimal programs, with the golden-rule equilibrium (associated with a program yielding maximum sustainable utility) replacing the notion of the von Neumann equilibrium.

With a shift in emphasis of many economies away from planning at the national level, there was a corresponding change in interpretation of dynamic optimization problems of the Ramsey type. The problem being solved was previously viewed as a normative problem the "social planner" ought to solve; it was now viewed as a descriptive problem that a typical representative agent (more precisely, an infinitely-lived dynasty of the typical agent) solves. The Ramsey objection to discounting future utilities as "ethically indefensible" on the part of the social planner was no longer relevant. If the representative agent did discount the future, the optimization problem would have to reflect this. Thus, the central problem to be solved in describing the agent's behavior would be a discounted dynamic optimization problem of the Ramsey-type.

This reformulation of the focus of the subject had two significant consequences. The issue of the existence of an optimal program, which had occupied center-stage for undiscounted dynamic optimization models, became a relatively unimportant aspect of the theory for discounted models, since it was a relatively straightforward exercise, under discounting, to establish the existence of an optimal program. In contrast, description of dynamic behavior of optimal programs became considerably more difficult.

Examples due to Kurz (1968) for continuous-time models, and Sutherland (1970) for discrete-time models indicated that there could be multiple stationary optimal states. Further, even if the stationary optimal state was unique, optimal programs starting from other initial states need not converge to it over time. Then, Samuelson (1973) presented an example due to Weitzman, which showed that optimal programs could cycle around a unique stationary state independent of the magnitude of the discount factor, and these cycles were not "boundary phenomena". While this destroyed any hope of a general turnpike theorem for discounted models, Samuelson conjectured that with (differential) strict concavity of the utility function, a turnpike property for optimal programs would continue to hold for high discount factors. This led to a considerable literature on the discounted turnpike problem. In alternative frameworks, Samuelson's conjecture was shown to be valid by Brock and Scheinkman (1976), Cass and Shell (1976), Rockafellar (1976) and Scheinkman (1976). Building on these papers, which appeared in a symposium of the Journal of Economic Theory, the theory was refined in the contributions of Araujo and Scheinkman (1977), Bewley (1980) and McKenzie (1982,1983), among others. A definitive survey of these developments are presented in McKenzie (1986).

A natural question that arose from this literature was how to describe dynamic optimal behavior when the discount factor was not close to unity. We will not go into the details of the history of how this question was tackled. It suffices to note that Chapters 3 - 13 of this book are devoted to various facets of this question, demonstrating the possibility and robustness of cyclical and chaotic optimal behavior in a variety of discounted dynamic optimization exercises. It should also be clear from these chapters that some important aspects of this question still remain unanswered.

We turn now to an overview of Sections 2-9 of this chapter. There are two aspects of the review of dynamic optimization theory that should be pointed out at this stage. First, to provide an introduction to the key ideas of this theory, the setting for this review is deliberately one-dimensional (the state space is a compact subset of the real line) although many of the topics covered here have their multidimensional counterparts in the literature. We provide some guidance to this literature in the bibliographic notes of Section 10. Second, we limit our coverage of dynamic optimal behavior to turnpike theory for high discount factors, leaving to the reader the task of sampling Chapters 3-13 of the book for the nature of dynamic optimal behavior when the discount factor is not close to unity.

Section 2 presents the standard reduced-form dynamic optimization model, as well as five examples of dynamic optimization exercises in economics which can be reduced to this standard form. The existence of optimal programs in this standard reduced form is given a complete self-contained treatment.

Two well-known characterizations of optimality have figured prominently in the literature of dynamic optimization. The first uses a primal approach, and characterizes optimality in terms of the existence a value function satisfying the functional equation of dynamic programming (often referred to as Bellman's optimality principle). This approach is developed in Section 3, where basic properties of the value and policy functions are derived.

The second characterization (usually referred to as the price characterization of optimality) is based on a dual approach, and is developed for convex structures, where separation theorems for convex sets play a crucial role. Here an optimal program is characterized in terms of existence of a sequence of dual variables or shadow prices, in terms of which (generalized) profit is maximized at the program at each date compared to any alternative activity available at that date and, in addition, an appropriate transversality condition is satisfied asymptotically. This material is developed in Section 4, where we establish the basic price characterization result , and study some of its implications.

Section 5 is devoted to sensitivity analysis (also known as comparative dynamics), where we examine how optimal policies (and therefore programs) change with changes in the parameters of the dynamic optimization model. Two aspects of such changes are studied in some detail : continuity (Section 5.1) and monotonicity (Section 5.2, 5.3). The former topic is based on the concept of uniform continuity, and the latter on the concept of supermodularity.

Section 6 presents the main result on the existence of a (non-trivial) stationary optimal stock, using the Kakutani fixed point theorem. The related concepts of a discounted golden-rule stock and a modified golden-rule are also introduced and discussed here.

While Sections 2-6 are developed using the methods of real and convex analysis, with differentiability assumptions playing no role, Sections 7-9 use the assumption of smooth preferences. Section 7 develops two basic properties of the dynamic optimization models under appropriate smoothness assumptions. First, if the utility function is continuously differentiable (in the interior of the transition possibility set), then the value function is continuously differentiable (in the interior of the state space), and the derivative of the value function is equal to the derivative of the utility function with respect to the initial state at the optimal point. Second, if the utility function is twice continuously differentiable (in the interior of the transition possibility set), then the policy function is differentiable at the stationary optimal stock, and the magnitude of the derivative of the policy function at the stationary optimal stock is equal to smaller (in absolute value) of the characteristic roots associated with the Ramsey-Euler equation at the stationary optimal stock.

Section 8 examines the issue of uniqueness of the stationary optimal stock. A simple example of non-uniqueness is constructed to show how easily this phenomenon can arise. Then, it is shown that there is a bound on the discount rate (depending on the transition possibility set and the utility function) such that for all higher discount rates, there is a unique (non-trivial) stationary optimal stock.

Section 9 is concerned with global asymptotic stability of the non-trivial stationary optimal stock. First, we present the example of Weitzman, as reported in Samuelson (1973), to show that even if the stationary optimal stock is unique, it is not globally asymptotically stable. Then, we pursue Samuelson's idea that if the Hessian of the utility function is negative definite at the golden-rule, then there is a bound on the discount factor (depending on the transition possibility set and the utility function), such that for all higher discount factors, there can be no cyclical optimal programs. This, in turn, yields global asymptotic stability of the non-trivial stationary optimal stock in our one-dimensional state space framework.

2. Dynamic Optimization Problems

We will describe the reduced-form dynamic optimization problem in Section 2.1 below. This is by now the standard form used to describe dynamic optimzation problems arising in economics. The reduced-form model is widely used because it has essentially a simple mathematical structure; one keeps track of the transition over time of only the state variable, from one state to another. It is also an extremely flexible model to interpret, and a wide variety of dynamic optimization problems in economics can be "reduced" to this form. We illustrate this by discussing five examples in Section 2.2.

2.1 The Reduced Form Model

A *state space* $X = [0, b] \subset \Re$ is given, where $b > 0$. Time is measured in discrete periods $t \in \{0, 1, 2, ...\} \equiv \mathbf{N}$. At each time t the state of the economic system is described by a number $x_t \in X$.

The typical dynamic optimization problem we will be concerned with seeks to maximize the objective functional

$$\sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) \tag{2.1}$$

over the set of all sequences $(x_t)_0^\infty$ satisfying the constraints:

$$(x_t, x_{t+1}) \in \Omega \qquad t \in \{0, 1, 2, ...,\} \tag{2.2}$$

$$x_0 = x \tag{2.3}$$

In the description of the above problem, δ is to be interpreted as a *discount factor*, u as a (reduced form) *utility function*, Ω as the *transition possiblity set*, and $x \in X$ as the *initial state.*

The following basic assumptions will be maintained throughout:

A1: $\Omega \subset X \times X$ *is a closed and convex set containing* $(0, 0)$.
A2: $u : \Omega \to \Re$ *is a bounded, concave, and upper semicontinuous function.*
A3: $\delta \in (0, 1)$.

Given (Ω, u, δ) and $x \in X$, we refer to a sequence $(x_t)_0^\infty$ satisfying (2.2) and (2.3) as a *program* (starting) from (the initial state) x. A solution to problem (2.1)-(2.3) is referred to as an *optimal program* from x.

Proposition 2.1: *If* (Ω, u, δ) *satisfies* A1-A3, *then there exists a solution to problem* (2.1)-(2.3) *for every initial state* $x \in X$.

Proof: Since u is bounded, we can find $B > 0$ such that $|u(x, y)| \leq B$ for all $(x, y) \in \Omega$. Thus, for every sequence $(x_t)_0^\infty$ satisfying (2.2),

$$\sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) \tag{2.4}$$

is an absolutely convergent series, and hence a convergent series, with

$$\sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) \leq B/(1 - \delta) \tag{2.5}$$

Let $S(x) = \sup\{\sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) : (x_t)_0^\infty$ is a sequence satisfying (2.2) and (2.3)$\}$. We will now show that this supremum is actually attained by some $(x_t)_0^\infty$ satisfying (2.2) and (2.3).

By definition of $S(x)$, there is a sequence of programs $(x_t^n)_{t=0}^\infty$, $n = 1, 2, 3, ...$, starting from the initial state x, such that

$$\sum_{t=0}^{\infty} \delta^t u(x_t^n, x_{t+1}^n) \geq S(x) - (1/n) \tag{2.6}$$

By the Cantor diagonal process, we can find a subsequence n' (of n) such that for each $t \in \{0, 1, 2, ...\}$, we have a number x_t^0, such that

$$x_t^{n'} \to x_t^0 \qquad \text{as } n' \to \infty \tag{2.7}$$

Since Ω is closed, $(x_t^0, x_{t+1}^0) \in \Omega$ for $t \in \{0, 1, 2, ...\}$, and $x_0^0 = x_0^{n'} = x$. Thus, $(x_t^0)_{t=0}^{\infty}$ is a program starting from x.

We claim that

$$\sum_{t=0}^{\infty} \delta^t u(x_t^0, x_{t+1}^0) = S(x). \tag{2.8}$$

If the claim were not true, then we could find $\varepsilon > 0$, such that

$$\sum_{t=0}^{\infty} \delta^t u(x_t^0, x_{t+1}^0) \leq S(x) - \varepsilon \tag{2.9}$$

Pick T large enough so that $B\delta^{T+1}/(1-\delta) \leq (\varepsilon/4)$. For $t \in \{0, 1, ..., T\}$, we can use (2.7) and the upper semicontinuity of u to get

$$\limsup_{n' \to \infty} u(x_t^{n'}, x_{t+1}^{n'}) \leq u(x_t^0, x_{t+1}^0) \tag{2.10}$$

We can now pick an interger $N > (4/\varepsilon)$, such that for $t \in \{0, 1, ..., T\}$, we have

$$u(x_t^{n'}, x_{t+1}^{n'}) \leq u(x_t^0, x_{t+1}^0) + [\varepsilon(1-\delta)/4]$$

whenever $n' \geq N$. Then, for $n' \geq N$, we can obtain the following string of inequalities:

$$\begin{aligned}
\sum_{t=0}^{\infty} \delta^t u(x_t^0, x_{t+1}^0) &\geq \sum_{t=0}^{T} \delta^t u(x_t^0, x_{t+1}^0) - (\varepsilon/4) \\
&\geq \sum_{t=0}^{T} \delta^t u(x_t^{n'}, x_{t+1}^{n'}) - (\varepsilon/2) \\
&\geq \sum_{t=0}^{\infty} \delta^t u(x_t^{n'}, x_{t+1}^{n'}) - 3(\varepsilon/4) \\
&\geq S(x) - (1/n') - 3(\varepsilon/4) \\
&> S(x) - \varepsilon
\end{aligned}$$

which contradicts (2.9) and establishes our claim (2.8). This means, of course, that $(x_t^0)_0^\infty$ is an optimal solution to (2.1)-(2.3). Q.E.D.

2.2 Examples

We now discuss five examples, arising in various contexts of capital theory, which can be treated as special cases of the reduced-form model of Section 2.1. For each example, described in terms of its "primitives", we show how it can be converted to its reduced form.

Example 2.1. (The One-Sector Model of Optimal Growth)

Let $G : \Re_+^2 \to \Re_+$ be a concave, non-decreasing, continuous and constant returns to scale *production function*, indicating the net output level $G(K, L)$ producible with inputs of capital (K) and labor (L). Labor is assumed to grow exogenously at a rate $n \geq 0$; that is $L_t = L_0(1+n)^t$ for $t \in \mathbf{N}$ and $L_0 > 0$. Capital depreciates at a constant rate d where $0 < d \leq 1$. The basic growth equation is

$$G(X_t, L_t) = C_{t+1} + K_{t+1} - (1-d)K_t$$

where $K_t \geq 0$ is the capital stock at time t, and $C_{t+1} \geq 0$ is the consumption at time $(t+1)$. Dividing through by L_{t+1}, and denoting $[K_t/L_t]$ by x_t and (C_{t+1}/L_{t+1}) by c_{t+1} for $t \in N$, we have

$$[G(x_t, 1)/(1+n)] = c_{t+1} + x_{t+1} - [(1-d)/(1+n)]x_t$$

Then letting $g(x) = G(x, 1)/(1+n)$ for $x \geq 0$, and $f(x) = g(x) + [(1-d)/(1+n)]x$ for $x \geq 0$, we obtain

$$c_{t+1} = f(x_t) - x_{t+1} \text{ for } t \in \mathbf{N}$$

A continuous, concave and increasing *welfare function*, $w : \Re_+ \to \Re$, evaluates the (period) welfare, $w(c)$, corresponding to a per capita consumption level, $c \geq 0$. The *discount factor*, $\delta \in (0, 1)$, indicates the weight placed by the planner (representative agent) on tomorrow's welfare relative to today's welfare in the intertemporal objective function.

The optimal growth problem can be written as:

$$(P1)\begin{cases} \text{Max} \sum_{t=0}^{\infty} \delta^t w(c_{t+1}) \\ \text{subject to } c_{t+1} = f(x_t) - x_{t+1}, \text{ and } (x_t, c_{t+1}) \geq 0 \\ \text{for } t \in \mathbf{N} \\ x_0 = x > 0 \end{cases}$$

We now indicate how to convert this example to its reduced form. First, we have to choose an appropriate state space. Let us assume that $\lim_{x\to\infty} [g(x)/x] = 0$. Then, we can find a number $b > 0$, such that $f(x) < x$ for $x > b$. In this case, if the initial capital-labor ratio, x_0, is in $[0, b]$, then all future capital-labor ratios, x_{t+1} (for $t \in \mathbf{N}$) must lie in $[0, b]$. Further, it is easy to verify that a solution to (P1) must have $x_t \in [0, b]$ for some finite t. Thus, it is legitimate to define the state space to be $X \equiv [0, b]$, since all the long-run dynamics is confined to this set.

Now, defining $\Omega = \{(x, z) \in \Re^2_+ : z \leq f(x)\}$, and $u : \Omega \to \Re$ by $u(x, z) = w(f(x) - z)$, a solution to problem (2.1)-(2.3) for (Ω, u, δ) corresponds precisely to a solution to problem (P1), and vice versa.

Example 2.2 (Two-Sector Model of Economic Growth)

The two-sector model of optimal economic growth, originally discussed by Uzawa (1964) and Srinivasan (1964), is a generalization of the one-sector model, discussed in Example 2.2 above. The severe restriction imposed in the one-sector model, that consumption can be traded off against investment on a one-to-one basis, is relaxed in the two-sector model, and it is principally this aspect that makes the two-sector model a considerably richer framework in studying economic growth problems than its one-sector predecessor.

Production uses two inputs, capital and labor. Given the total amounts of capital and labor available to the economy, the inputs are allocated to two "sectors" of production, the consumption good sector, and the investment good sector. Output of the former sector is consumed (and cannot be used for investment purposes), and output of the latter sector is used to augment the capital stock of the economy (and cannot be consumed). Thus, the consumption-investment decision amounts to a decision regarding allocation of capital and labor between the two production sectors.

The discounted sum of outputs of the consumption good sector is to be maximized to arrive at the appropriate sectoral allocation of inputs in each period. [A welfare function on the output of the consumption good sector can be used, instead of the output of the consumption good sector itself, in the objective function, but it is usual to assimilate this welfare function, when it is increasing and concave, in the "production function" of the consumption good sector.]

Formally, the model is specified by (F, G, d, δ), where

(a) the *production function* in the *consumption good sector*, $F : \Re^2_+ \to \Re_+$, satisfies:

(i) F is continuous and homogeneous of degree one on $\Re^2_+$.

(ii) F is non-decreasing on $\Re^2_+$.

(iii) F is concave on $\Re^2_+$.

(b) the *production function* in the *investment good sector*, $G : \Re^2_+ \to \Re_+$, satisfies:

(i) G is continuous and homogeneous of degree one on $\Re^2_+$.

(ii) G is non-decreasing on $\Re^2_+$.

(iii) G is concave on $\Re^2_+$.

(iv) $\lim_{K\to\infty} [G(K,1)/K] = 0$

(c) the *depreciation factor*, d, satisfies $0 < d \leq 1$;

(d) the *discount factor*, δ, satisfies $0 < \delta < 1$.

The optimal growth problem can be written as:

$$(P2)\left\{\begin{array}{c} \text{Max } \sum_{t=0}^{\infty} \delta^t c_{t+1} \\ \text{subject to } c_{t+1} = F(k_t, n_t) \text{ for } t \in \mathbf{N} \\ x_{t+1} = G(x_t - k_t, 1 - n_t) + (1-d)x_t \text{ for } t \in \mathbf{N} \\ 0 \leq k_t \leq x_t,\ 0 \leq n_t \leq 1, \text{ for } t \in \mathbf{N} \\ x_0 = x > 0 \end{array}\right.$$

Here x_t is the total capital available at date t, which is allocated between the consumption good sector (k_t) and the investment good sector ($x_t - k_t$). Labor is exogenously available at a constant amount (normalized to unity), which is allocated between the consumption good sector (n_t) and the investment good sector ($1 - n_t$). Note that an exogenously growing labor force (at a constant growth rate) can be accommodated easily by interpreting k_t and x_t as per worker capital stocks, and reinterpreting the depreciation factor, d.

As in the one-sector model, we can find $b > 0$, such that for $x \in [0, b]$, we have $G(x, 1) + (1 - d)x$ in $[0, b]$, and furthermore for any solution to (P2), $x_t \in [0, b]$ for some t. Thus, it is appropriate to define $X = [0, b]$ as the state space.

To convert the problem to its reduced-form, we can define the transition possibility set, Ω, as:

$$\Omega = \{(x, z) \text{ in } \Re_+^2 : z \leq G(x, 1) + (1 - d)x\}$$

and the utility function, u, as

$$\begin{aligned} u(x, z) &= \text{Max } \; F(k, n) \\ \text{Subject to } 0 &\leq k \leq x, \; 0 \leq n \leq 1 \\ z &\leq G(x - k, 1 - n) + (1 - d)x \end{aligned}$$

Then a solution to (P2) corresponds exactly to a solution (2.1)-(2.3) for (Ω, u, δ), and vice versa.

Example 2.3 (Optimal Exploitation of a Fishery)

The economic exploitation of a fishery figures prominently in the literature on the management of renewable resources (see, for example, Clark (1976) for references to the relevant literature). The following is an idealized description (following Dasgupta (1982)) of the relevant optimization problem.

The resource stock (biomass of the fish species) at the end of period t is denoted by x_t. The biological reproduction function (also called the stock recruitment function) yields the biomass, y_{t+1}, in period $(t + 1)$ given by

$$y_{t+1} = f(x_t)$$

If c_{t+1} is harvested in period $(t+1)$ then the stock at the end of time period $(t+1)$ is x_{t+1} given by

$$x_{t+1} = f(x_t) - c_{t+1}$$

It is usual to assume that $f(0) = 0$, f is continuous, non-decreasing and concave, with $\lim_{x \to 0}[f(x)/x] > 1$ and $\lim_{x \to \infty}[f(x)/x] = 0$ [although there are important variations, especially dealing with non-concave functions, f].

We can find $b > 0$ such that whenever x_t is in $[0, b]$, x_{t+1} is in $[0, b]$ also. Thus, $[0, b]$ is a legitimate choice for the state space, X.

Harvesting the fishery is not costless. Specifically, the harvest, c_{t+1}, depends on the resource biomass available for exploitation, y_{t+1}, and the labor effort, e_{t+1}:

$$c_{t+1} = G(y_{t+1}, e_{t+1})$$

with G a non-decreasing, concave function in both its variables (increasing in both arguments when (y, e) is stricty positive), such that given any $y \geq 0$, $G(y, e) \leq y$ for all e. We express the effort, e_{t+1}, required to havest c_{t+1} when the biomass available for harvest is y_{t+1} as:

$$e_{t+1} = H(c_{t+1}, y_{t+1})$$

with H increasing in its first argument and decreasing in its second. This is the cost of exploiting the fishery. The benefit obtained depends on the harvest level, c_{t+1}:

$$B_{t+1} = B(c_{t+1})$$

where B is an increasing, concave function. [For a competitive fishery, B is linear, with $B(c_{t+1}) = pc_{t+1}$, where $p > 0$ is the price (assumed constant over time) per unit of the fish.] The return in period $(t+1)$ is then the benefit minus the cost:

$$w(c_{t+1}, x_t) = B(c_{t+1}) - H(c_{t+1}, f(x_t))$$

The optimization problem can then be written as

$$(P3)\begin{cases} \text{Max} \sum_{t=0}^{\infty} \delta^t w(c_{t+1}, x_t) \\ \text{Subject to } c_{t+1} = f(x_t) - x_{t+1} \text{ for } t \in \mathbf{N} \\ c_{t+1} \geq 0, x_{t+1} \geq 0 \text{ for } t \in \mathbf{N} \\ x_0 = x > 0 \end{cases}$$

To convert the problem to its reduced form, we define the transition possibility set, Ω, as:

$$\Omega = \{(x, z) \text{ in } \Re_+^2 : z \leq f(x)\}$$

and the utility function, u, as:

$$u(x, z) = w(f(x) - z, x)$$

Then, a solution to (P3) corresponds exactly to a solution to (2.1)-(2.3) for (Ω, u, δ).

Example 2.4 (A Three-Sector Model with an Aging Process)

This example is due to Weitzman, as reported in Samuelson (1973). It involves a pure aging process in one of the sectors, which produces wine, and therefore might be called an "Austrian" capital-theoretic model.

The representative agent has one unit of labor. She can allocate it between two sectors: one of these sectors produces bread, the other produces grape juice. It takes one unit of labor per unit of bread production, and one unit of labor per unit of grape juice production. A third sector produces wine, and its production process requires one unit of grape juice per unit of wine, and one period of time (for the juice to ferment).

As a consumer, the agent derives satisfaction from the consumption of bread and wine, but not from grape juice.

Let x_t denote the allocation of labor in period t to grape-juice production, with $x_0 > 0$ (initial allocation) as given. Then $(1-x_t)$ of labor is allocated to production of bread. Thus, in period $(t+1)$, where $t \geq 0$, we have bread production of $(1 - x_{t+1})$, grape-juice production of x_{t+1}, and wine produciton of x_t.

Assume that the consumer's felicity function is

$$f(b, w) = b^\alpha w^\beta, \text{ with } (\alpha, \beta) >> 0, \ \alpha + \beta \leq 1$$

where b and w are the amounts of bread and wine consumed.

The agent's optimization problem can be written as:

$$(P4) \begin{cases} \text{Max } \sum_{t=0}^{\infty} \delta^t f(b_{t+1}, w_{t+1}) \\ \text{Subject to } b_{t+1} = 1 - x_{t+1} \text{ for } t \in \mathbf{N} \\ 0 \leq x_t \leq 1 \text{ for } t \in \mathbf{N} \\ x_0 = x > 0 \end{cases}$$

To convert the problem to its reduced form, we can define the state space $X = [0, 1]$, the transition possibility set, Ω, to be:

$$\Omega = \{(x, z) \in X^2\}$$

and the utility function, u, to be:

$$u(x, z) = f(1 - x_{t+1}, x_t) = (1 - x_{t+1})^\alpha x_t^\beta$$

Then an optimal solution to (P4) corresponds exactly to an optimal solution to (2.1)-(2.3) for (Ω, u, δ).

Example 2.5 (A Model of Forest Management)

This example is a simplified version of the forest management model, developed in Mitra and Wan (1985, 1986), based on the pioneering work in this area by Faustmann (1849). Our exposition of the example is based on the presentation of Wan (1987).

An agent has a plot of land (normalized to unity), which is good for growing a certain type of tree. The growth process for the tree is as follows: after saplings are planted, the tree grows in timber content for two years, the timber content being a $\in (0, 1)$ after one year, and 1 after two years (per unit of land). After two years, the tree decays and the timber becomes worthless.

The agent derives a return from selling the timber, the timber being assumed of homogeneous quality, whether it is from one or two year old trees. Planting and cutting trees are assumed costless activities.

Let x_t denote the amount of land with two-year old trees at the end of period t. Then, $(1-x_t)$ is the amount of land with one-year old trees at the end of period t.

At the end of period t, all two year old trees are cut down and sold as timber. This clears up x_t of land for reforestation. In contrast one-year old trees might or might not be cut down. The one-year old trees which are *not* cut down at the end of period t become the two-year old trees at the end of period $(t+1)$. Thus, the land cleared up by cutting down one-year old trees at the end of period t is $[(1-x_t)-x_{t+1}]$. The total timber content of all trees cut down at the end of period t is, therefore, given by

$$c_t = x_t + a[(1-x_t) - x_{t+1}]$$

The agent's optimization problem can be written as:

$$(P5)\begin{cases} \text{Max } \sum_{t=0}^{\infty} \delta^t R(c_t) \\ \text{Subject to } c_t = x_t + a[(1-x_t) - x_{t+1}] \text{ for } t \in \mathbf{N} \\ x_{t+1} \leq (1-x_t) \text{ for } t \in \mathbf{N} \\ 0 \leq x_t \leq 1, \text{ for } t \in \mathbf{N} \\ x_0 = x \geq 0 \end{cases}$$

To convert this problem to its reduced form, we can define the state space to be $X = [0,1]$, the transition possibility set, Ω, to be:

$$\Omega = \{(x,z) \in X^2 : z \leq (1-x)\}$$

and the utility function, u, to be:

$$u(x,z) = R(a + (1-a)x - az)$$

Then, an optimal solution to (P5) corresponds to an optimal solution to (2.1)-(2.3) for (Ω, u, δ).

3. Dynamic Programming

Dynamic optimization problems of the type described in Section 2 can be studied conveniently by the method of dynamic programming. In this section, we provide the basic results that can be established for the class of problems defined by (2.1)-(2.3) by using the dynamic programming approach.

We can associate with each dynamic optimization problem two functions, called the (optimal) *value function* and the (optimal) *policy function* [or, more generally, the (optimal) policy corresondence]. The value function defines the maximized value of the objective function (2.1), given the constraints (2.2), (2.3), corresponding to each initial state, x. If there is a unique optimal solution $(x_t^*)_0^\infty$ to the problem (2.1)-(2.3) for each initial state $x \in X$, the policy function defines the first period optimal state, x_1^*, corresponding to each initial state, x. [More generally, the policy correspondence describes the set of states that it is optimal to go to in period 1, corresponding to each initial state, x.]

Given the stationary recursive nature of the optimization problem, the policy function, in fact, generates the entire optimal solution, starting from any initial state. The value function helps us to study the properties of the policy function, given the connection between the two through the functional equation of dynamic programming. Thus, the value and policy functions are extremely useful objects of study in describing dynamic optimal behavior.

3.1 The Value Function

Given the existence of an optimal solution to problem (2.1)-(2.3) fore every $x \in X$ [Proposition 2.1], we can define the *value function*, $V : X \to \Re$ as follows:

$$V(x) = \sum_{t=0}^{\infty} \delta^t u(x_t^*, x_{t+1}^*) \tag{3.1}$$

where $(x_t^*)_0^\infty$ is an optimal solution to (2.1)-(2.3) corresponding to the initial state $x \in X$.

The following result summarizes the basic properties of the value function.

Proposition 3.1: (i) *The value function, V, is a concave and continuous function on X.* (ii) *V satisfies the following functional equation of dynamic programming*

$$V(x) = \max_{y \in \Omega_x}\{u(x,y) + \delta V(y)\} \tag{3.2}$$

for all $x \in X$ [where $\Omega_x = \{y \in X : (x,y) \in \Omega\}$]. (iii) *$V$ is the only continuous function on X which satisfies (3.2).* (iv) *$(x_t)_0^\infty$ is an optimal solution to (2.1)-(2.3) if and only if*

$$V(x_t) = u(x_t, x_{t+1}) + \delta V(x_{t+1}) \text{ for } t \in \mathbf{N} \tag{3.3}$$

Proof: (i) Let x, $x' \in X$ and let $0 < \lambda < 1$. Let $(x_t)_0^\infty$ and $(x'_t)_0^\infty$ be optimal programs from x and x' respectively. Define $x'' = \lambda x + (1-\lambda)x'$, and let (x''_t) be an optimal program from x''. By convexity of Ω, the sequence $(\lambda x_t + (1-\lambda)x'_t)_0^\infty$ is a program from x''. Thus, we have

$$\begin{aligned}
V(x'') &= \sum_{t=0}^{\infty} \delta^t u(x''_t, x''_{t+1}) \\
&\geq \sum_{t=0}^{\infty} \delta^t u(\lambda x_t + (1-\lambda)x'_t, \lambda x_{t+1} + (1-\lambda)x'_{t+1}) \\
&\geq \sum_{t=0}^{\infty} \delta^t [\lambda u(x_t, x_{t+1}) + (1-\lambda)u(x'_t, x'_{t+1})] \\
&= \lambda V(x) + (1-\lambda)V(x')
\end{aligned}$$

the second inequality following from the concavity of u on Ω. Thus, V is concave on X.

In order to establish the continuity of V on X, we first establish its upper semicontinuity. This can be done by following essentially the method used in the proof of Proposition 2.1.

If V were not upper semicontinuous on X, we can find $x^n \in X$ for $n = 1, 2, 3, \ldots$, with $x^n \to x^0$ and $V(x^n) \to V$ as $n \to \infty$, with $V > V(x^0)$. Let us denote by (x_t^n) an optimal program from x^n for $n = 0, 1, 2, \ldots$.

Denoting $[V - V(x^0)]$ by ε, we can find T large enough so that $B\delta^{T+1}/(1-\delta) \leq (\varepsilon/4)$.

Clearly, we can find a subsequence n' (of n) such that for $t \in \{0, 1, ..., T\}$

$$x_t^{n'} \to x_t^0 \text{ as } n' \to \infty.$$

Then, using the upper semicontinuity of u, we can find N such that for $t \in \{0, 1, ..., T\}$

$$u(x_t^{n'}, x_{t+1}^{n'}) \leq u(x_t^0, x_{t+1}^0) + [\varepsilon(1-\delta)/4]$$

whenever $n' \geq N$. Thus, for $n' \geq N$, we obtain the following string of inequalities

$$\begin{aligned} V(x^0) &= \sum_{t=0}^{\infty} \delta^t u(x_t^0, x_{t+1}^0) \\ &\geq \sum_{t=0}^{T} \delta^t u(x_t^0, x_{t+1}^0) - (\varepsilon/4) \\ &\geq \sum_{t=0}^{T} \delta^t u(x_t^{n'}, x_{t+1}^{n'}) - (\varepsilon/2) \\ &\geq \sum_{t=0}^{\infty} \delta^t u(x_t^{n'}, x_{t+1}^{n'}) - 3(\varepsilon/4) \\ &= V(x^{n'}) - 3(\varepsilon/4) \end{aligned}$$

Since $V(x^{n'}) \to V$, we have $V(x^0) \geq V - 3(\varepsilon/4) > V - \varepsilon = V(x^0)$, a contradiction. Thus, V is upper semicontinuous on X.

V is concave on X, and hence continuous on *int* $X = (0, b)$. To show that V is continuous at 0, let x^n be a sequence of points in X ($n = 1, 2, 3, ...$) converging to 0. Then $V(x^n) = V((1 - (x^n/b))\ 0 + (x^n/b)b) \geq [1 - (x^n/b)]V(0) + (x^n/b)V(b)$. Letting $n \to \infty$, $\liminf_{n\to\infty} V(x^n) \geq V(0)$. On the other hand, since V is upper semicontinuous on X, $\limsup_{n\to\infty} V(x^n) \leq V(0)$. Thus, $\lim_{n\to\infty} V(x^n)$ exists and equals $V(0)$. The continuity of V at b is established similarly.

(ii) Let $y \in \Omega_x$, and let $(y_t)_0^\infty$ be an optimal program from y. Then $(x, y_0, y_1, ...)$ is a program from x, and hence, by definition of V,

$$\begin{aligned} V(x) &\geq u(x, y_0) + \sum_{t=1}^{\infty} \delta^t u(y_{t-1}, y_t) \\ &= u(x, y_0) + \delta \sum_{t=1}^{\infty} \delta^{t-1} u(y_{t-1}, y_t) \\ &= u(x, y_0) + \delta \sum_{t=0}^{\infty} \delta^t u(y_t, y_{t+1}) \\ &= u(x, y_0) + \delta V(y) \end{aligned}$$

So, we have established that

$$V(x) \geq u(x, y) + \delta V(y) \text{ for all } y \in \Omega_x \tag{3.4}$$

Next, let $(x_t)_0^\infty$ be an optimal program from x, and note that

$$\begin{aligned} V(x) &= u(x_0, x_1) + \delta \left[\sum_{t=1}^{\infty} \delta^{t-1} u(x_t, x_{t+1}) \right] \\ &= u(x_0, x_1) + \delta \left[\sum_{t=0}^{\infty} \delta^t u(x_{t+1}, x_{t+2}) \right] \\ &\leq u(x_0, x_1) + \delta V(x_1) \end{aligned}$$

Using (3.4), we then have

$$V(x) = u(x_0, x_1) + \delta V(x_1) \tag{3.5}$$

Now, (3.4) and (3.5) establish (3.2).

(iii) Let W be a continuous function on X satisfying (3.2). Let $D \equiv \max_{x \in X} |V(x) - W(x)|$. Given any $x \in X$, denote by y a solution to the maximization problem on the right-hand side of (3.2). Then, $V(x) = u(x, y) + \delta V(y)$, and $W(x) \geq u(x, y) + \delta W(y)$, so that

$$[V(x) - W(x)] \leq \delta[V(y) - W(y)] \leq \delta D \tag{3.6}$$

Denote by y' a solution to the maximization problem on the right-hand side of (3.2), when V is replaced by W. Then, $W(x) = u(x, y') + \delta W(y')$, and $V(x) \geq u(x, y') + \delta V(y')$, so that

$$[W(x) - V(x)] \leq \delta[W(y') - V(y')] \leq \delta D \tag{3.7}$$

Using (3.6) and (3.7) we get

$$|V(x) - W(x)| \leq \delta D$$

and since $x \in X$ was arbitrary,

$$D \equiv \max_{x \in X} |V(x) - W(x)| \leq \delta D$$

Thus, $D = 0$, and so $W = V$ on X.

(iv) If $(x_t)_0^\infty$ satisfies (3.3), then we get for any $T \geq 1$,

$$V(x) = \sum_{t=0}^{T} \delta^t u(x_t, x_{t+1}) + \delta^{T+1} V(x_{T+1})$$

Since $|V(x)| \leq B/(1 - \delta)$ for all $x \in X$ and $\delta^{T+1} \to 0$ as $T \to \infty$, we have

$$V(x) = \sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1})$$

Then, by the definition of V, $(x_t)_0^\infty$ is an optimal program (from x).

To establish the converse implication, let $(x_t^*)_0^\infty$ be an optimal solution to (2.1)-(2.3). Then, for each $T \geq 1$, $(x_T^*, x_{T+1}^*, ...)$ is an optimal solution to (2.1)-(2.3) for $x = x_T^*$. For, if $(y_t)_0^\infty$ is a program from $y_0 = x_T^*$ such that

$$\sum_{t=0}^{\infty} \delta^t u(y_t, y_{t+1}) > \sum_{t=T}^{\infty} \delta^{t-T} u(x_t^*, x_{t+1}^*)$$

then the sequence $(x_0^*, ..., x_T^*, y_1, y_2, ...)$ has a discounted sum of utilities

$$\sum_{t=0}^{T-1} \delta^t u(x_t^*, x_{t+1}^*) + \sum_{t=T}^{\infty} \delta^t u(y_t, y_{t+1}) > \sum_{t=0}^{\infty} \delta^t u(x_t^*, x_{t+1}^*)$$

which contradicts the optimality of $(x_t^*)_0^\infty$ from $x = x_0^*$. Now, using the result (3.5) in (ii) above, we have $V(x_t^*) = u(x_t^*, x_{t+1}^*) + \delta V(x_{t+1}^*)$ for $t \in \mathbf{N}$.

3.2 The Policy Function

An optimal solution to the problem (2.1)-(2.3) need to be unique. For some states x_t, there might be more than one successor state x_{t+1} satisfying (3.3). To ensure that an optimal solution to (2.1)-(2.3) is unique for every $x \in X$, we use a strict concavity assumption on u.

A4: u is strictly concave in its second argument

If A4 holds, then given any $x \in X$, there is a unique solution to the maximization problem on the right-hand side of (3.2). For if y and $y'(y' \neq y)$ in Ω_x both solved this problem, then $V(x) = u(x,y) + \delta V(y)$ and $V(x) = u(x,y') + \delta V(y')$. However, $(x, 0.5y + 0.5y') \in \Omega$, and $u(x, 0.5y + 0.5y') + \delta V(0.5y + 0.5y') > 0.5u(x,y) + 0.5u(x,y') + \delta[0.5V(y) + 0.5V(y')] = V(x)$, a contradiction to (3.2).

For each x, denote the unique state $y \in \Omega_x$ which solves the maximization problem on the right hand side of (3.2) by $h(x)$. We will call $h : X \to X$ the *policy function.*

The following result summarizes the basic properties of the policy function.

Proposition 3.2: (i) *The policy function $h : X \to X$ is continuous on X.*

(ii) *For all $(x,y) \in \Omega$ with $y \neq h(x)$ we have*

$$u(x,y) + \delta V(y) < V(x) = u(x, h(x)) + \delta V(h(x)) \tag{3.8}$$

(iii) *$(x_t)_0^\infty$ is an optimal solution to (2.1)-(2.3) if and only if*

$$x_{t+1} = h(x_t) \text{ for } t \in \mathbf{N} \tag{3.9}$$

Proof: (i) Let $(x^n)(n = 1,2,3,...)$ be a sequence of points in X converging to x^0. Then for $n = 1,2,3,....$

$$V(x^n) = u(x^n, h(x^n)) + \delta V(h(x^n)) \tag{3.10}$$

Let $(y^{n'})$ be an arbitrary convergent subsequence of $(h(x^n))$ converging to y^0. We claim that $y^0 = h(x^0)$. For the subsequence n', using (3.10), the upper semicontinuity of u and the continuity of V,

$$\begin{aligned} V(x^0) &= \limsup_{n' \to \infty} u(x^{n'}, y^{n'}) + \delta V(y^0) \\ &\leq u(x^0, y^0) + \delta V(y^0) \end{aligned}$$

But, by (3.2), $V(x^0) \geq u(x^0, y^0) + \delta V(y^0)$, so that $V(x^0) = u(x^0, y^0) + \delta V(y^0)$. This means that $y^0 = h(x^0)$, establishing our claim. Thus, $h(x^n) \to h(x^0)$ as $n \to \infty$, establishing continuity of h.

(ii) We have $V(x) = u(x, h(x)) + \delta V(h(x))$ from (3.2) and the definition of h. For all $y \in \Omega_x$, we have $u(x, y) + \delta V(y) \leq V(x)$ by (3.2). Further, if equality holds in the previous weak inequality, then $y = h(x)$. Thus if $y \neq h(x)$, the strict inequality must hold. This establishes (3.8).

(iii) If $(x_t)_0^\infty$ is an optimal solution to (2.1)-(2.3), then by Proposition 3.1 (iv), for every $t \in \mathbf{N}$,

$$V(x_t) = u(x_t, x_{t+1}) + \delta V(x_{t+1})$$

so that by Proposition 3.2 (ii), $x_{t+1} = h(x_t)$.

Conversely, if (3.9) holds, then by Proposition 3.2 (ii), we have for $t \in \mathbf{N}$

$$V(x_t) = u(x_t, x_{t+1}) + \delta V(x_{t+1})$$

so that $(x_t)_0^\infty$ is an optimal solution to (2.1)-(2.3) by Proposition 3.1 (iv).

4. Duality Theory

Optimal programs can be characterized in terms of dual variables (or shadow prices). At the shadow prices supporting an optimal program, there is no activity which yields a higher "generalized profit" at any date (value of utility plus value of terminal stocks minus value of initial stocks at that date) than the activity chosen along the optimal program at that date. This support property is referred to as the *competitive condition*, and the shadow prices as *competitive prices.*

In addition, optimal programs must satisfy the condition that the aymptotic value of the stocks (at the shadow prices) must be zero, which is referred to as the *transversality condition.*

The two conditions taken together are also sufficient to ensure optimality, so that they provide a complete characterization of optimal programs, generally called a *price characterization of optimality.*

It was observed by Weitzman (1973) that the shadow prices appearing in the competitive condition also provide a support for the value function, which was introduced in Section 3. Given the competitive condition, the validty of the transversality condition is equivalent to the property that the competitive prices support the value function.

This connection between dynamic programming and duality theory forms the basis of Weitzman's approach to the price characterization of optimal programs, which we will describe below.

It is convenient (although not essential) to develop the theory under an additional "monotonicity assumption", which ensures that the relevant competitive prices are non-negative at each date.

A5: If $(x, z) \in \Omega$, and $x' \in X$, $x' \geq x$, $0 \leq z' \leq z$, then $(x', z') \in \Omega$ and $u(x', z') \geq u(x, z)$.

Proposition 4.1: *If $(x_t)_0^\infty$ is a program from $x \in X$, and there is a sequence $(p_t)_0^\infty$ such that $p_t \in \Re_+$ for $t \in \mathbf{N}$, and the following conditions hold:*

$$\text{(i) } \delta^t u(x_t, x_{t+1}) + p_{t+1}x_{t+1} - p_t x_t \geq \delta^t u(x, z) + p_{t+1}z - p_t x \quad \textit{for all } (x, z) \in \Omega,\ t \in \mathbf{N} \tag{4.1}$$

$$\text{(ii) } \lim_{t\to\infty} p_t x_t = 0 \tag{4.2}$$

then $(x_t)_0^\infty$ is an optimal program from x.

Proof: Let $(x'_t)_0^\infty$ be any program from x. Using (4.1), we have for $t \geq 0$,

$$\delta^t[u(x'_t, x'_{t+1}) - u(x_t, x_{t+1})] \leq (p_{t+1}x_{t+1} - p_t x_t) - (p_{t+1}x'_{t+1} - p'_t x'_t)$$

Summing this from $t = 0$ to $t = T$

$$\begin{aligned}\sum_{t=0}^{T} \delta^t[u(x'_t, x'_{t+1}) - u(x_t, x_{t+1})] &\leq (p_{T+1}x_{T+1} - p_0 x_0) - \\ &\quad (p_{T+1}x'_{T+1} - p_0 x'_0) \\ &= (p_{T+1}x_{T+1} - p_{T+1}x'_{T+1}) \\ &\leq p_{T+1}x_{T+1}\end{aligned}$$

Letting $T \to \infty$, we get

$$\sum_{t=0}^{\infty} \delta^t u(x'_t, x'_{t+1}) - \sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) \leq \lim_{T\to\infty} p_{T+1}x_{T+1} = 0$$

which proves that $(x_t)_0^\infty$ is optimal from x.

Remark 4.1: (i) Condition (4.2) can be replaced by

$$\liminf_{t\to\infty} p_t x_t = 0 \tag{4.3}$$

in Proposition 4.1.

(ii) Condition 4.2 is called the *transversality condition.* Its significance was first noted by Malinvaud (1953) in his study of intertemporal efficiency. Condition 4.1 is referred to as the *competitive condition*; the sequence $(p_t)_0^\infty$ of dual variables are called *competitive prices.* The result of Proposition 4.1 is then paraphrased as: "A competitive program satisfying the transversality condition is optimal."

In contrast to Proposition 4.1, the converse result (Proposition 4.2 below) exploits the convex structure of the dynamic optimization model.

Proposition 4.2: *Suppose* $(x_t)_0^\infty$ *is an optimal program from* $x > 0$. *Suppose also that there is some* $(\widehat{x}, \widehat{z})$ *in* Ω *with* $\widehat{z} > 0$.

Then, there is a sequence $(p_t)_0^\infty$ *such that* $p_t \in \Re_+$ *for* $t \in \mathbf{N}$ *and the following conditions hold:*

(i) $\delta^t u(x_t, x_{t+1}) + p_{t+1}x_{t+1} - p_t x_t \geq \delta^t u(x,z) + p_{t+1}z - p_t x$
for all $(x,z) \in \Omega$ *and* $t \in \mathbf{N}$ (4.4)

(ii) $\delta^t V(x_t) - p_t x_t \geq \delta^t V(x) - p_t x$
for all $x \in X$ *and* $t \in \mathbf{N}$ (4.5)

(iii) $\lim_{t\to\infty} p_t x_t = 0$ (4.6)

In order to establish Proposition 4.2, we first state and prove two Lemmas.

Lemma 4.1: *If* $x \in X$ *and* $x > 0$ *, then there is* $p \in \Re_+$ *such that*

$$V(x) - px \geq V(y) - py \quad \text{for all } y \in X \tag{4.7}$$

Proof: If $x \in (0,b]$, and we define p to be the left-hand derivative of V at x, then (4.7) holds. [Note that the left-hand derivative of V is well defined for all $x \in (0,b)$ by concavity of V, and also at $x = b$ by A5]. A5 ensures that p is non-negative.

Lemma 4.2: *Suppose* (x_t) *is an optimal program from* $\overline{x} \in X$, *and suppose also that there is* $(\widehat{x}, \widehat{z})$ *in* Ω *with* $\widehat{z} > 0$. *If for some* $t \in \mathbf{N}$ *there is* $p_t \in \Re_+$ *satisfying*

$$\delta^t V(x_t) - p_t x_t \geq \delta^t V(x) - p_t x \text{ for all } x \in X \tag{4.8}$$

then there is $p_{t+1} \in \Re_+$ *satisfying*

$$\delta^{t+1} V(x_{t+1}) - p_{t+1}x_{t+1} \geq \delta^{t+1} V(x) - p_{t+1}x \quad \text{for all } x \in X \tag{4.9}$$

and, furthermore,

$$\begin{aligned}\delta^t u(x_t, x_{t+1}) + p_{t+1}x_{t+1} - p_t x_t \;&\geq\; \delta^t u(x,y) \\ +p_{t+1}y - p_t x \text{ for all } (x,y) \;&\in\; \Omega\end{aligned} \tag{4.10}$$

Proof: Using Proposition 3.1, we have $V(x_t) = u(x_t, x_{t+1}) + \delta V(x_{t+1})$, and for all $(x, y) \in \Omega$, $V(x) \geq u(x, y) + \delta V(y)$. Using these in (4.8), we get

$$\begin{aligned} \theta_{t+1} &\equiv \delta^t u(x_t, x_{t+1}) + \delta^{t+1} V(x_{t+1}) - p_t x_t \\ &\geq \delta^t u(x, y) + \delta^{t+1} V(y) - p_t x \text{ for all } (x, y) \in \Omega \end{aligned}$$

Thus, we have

$$\theta_{t+1} - \delta^t u(x, y) + p_t x \geq \delta^{t+1} V(y) \text{ for all } (x, y) \in \Omega \tag{4.11}$$

Define two sets A and B as follows:

$$\begin{aligned} \text{A} &= \{(w, y) \in \Re^2 : (x, y) \in \Omega \text{ for some } x \in X \\ \text{and } w &> \theta_{t+1} - \delta^t u(x, y) + p_t x\} \end{aligned}$$

$$\text{B} = \{(w, y) \in \Re^2 : y \in X \text{ and } w \leq \delta^{t+1} V(y)\}$$

Clearly, A and B are non-empty and convex (since u is concave on Ω and V is concave on X). Also, by (4.11), A and B are disjoint. Hence, by the Minkowski separation theorem, there is (μ, v) in $\Re^2$ with $(\mu, v) \neq 0$ and $\alpha \in \Re$ such that

$$\mu w + v y \geq \alpha \text{ for all } (w, y) \in A \tag{4.12}$$

$$\mu w + v y \leq \alpha \text{ for all } (w, y) \in B \tag{4.13}$$

Using (4.13), we have $\mu \geq 0$. Define $q_{t+1} = (-v)$. Then, using (4.12) and (4.13),

$$\begin{aligned} \mu[\theta_{t+1} - \delta^t u(x, y) + p_t x] - q_{t+1} y &\geq \\ \mu \delta^{t+1} V(y') - q_{t+1} y' \text{ for all } (x, y) &\in \Omega \\ \text{and all } y' &\in X \end{aligned} \tag{4.14}$$

Substituting $x = x_t$ and $y = x_{t+1}$ in (4.14),

$$\begin{aligned} \mu \delta^{t+1} V(x_{t+1}) - q_{t+1} x_{t+1} &\geq \mu \delta^{t+1} V(y') - q_{t+1} y' \\ \text{for all } y' &\in X \end{aligned} \tag{4.15}$$

Substituting $y' = x_{t+1}$ in (4.14),

$$\begin{aligned} \mu[\delta^t u(x_t, x_{t+1}) - p_t x_t] + q_{t+1} x_{t+1} &\geq \\ \mu[\delta^t u(x, y) - p_t x] + q_{t+1} y \text{ for all } (x, y) &\in \Omega \end{aligned} \tag{4.16}$$

We claim now that $\mu \neq 0$. For if $\mu = 0$, then by (4.15), we have $q_{t+1}x_{t+1} \leq q_{t+1}y'$ for all y' in X, while by (4.16), $q_{t+1}x_{t+1} \geq q_{t+1}y$ for all y such that $(x, y) \in \Omega$ for some $x \in X$. Thus

$$q_{t+1}x_{t+1} = q_{t+1}y \text{ for all } y \text{ such that } (x, y) \in \Omega \text{ for some } x \in X \tag{4.17}$$

Since $(\widehat{x}, \widehat{z}) \in \Omega$ with $\widehat{z} > 0$ and $(\widehat{x}, 0) \in \Omega$ by A5, we have $q_{t+1}x_{t+1} = q_{t+1}\widehat{z} = q_{t+1}0 = 0$. Thus, $q_{t+1} = (-v) = 0$, and so $(\mu, v) = 0$, a contradiction. Thus, $\mu \neq 0$, and since $\mu \geq 0$, we have $\mu > 0$. Define $p_{t+1} = (q_{t+1}/\mu)$ and use (4.15), (4.16) to get

$$\begin{aligned} \delta^{t+1}V(x_{t+1}) - p_{t+1}x_{t+1} &\geq \delta^{t+1}V(y') - p_{t+1}y' \\ \text{for all } y' &\in X \end{aligned} \tag{4.18}$$

$$\begin{aligned} \delta^t u(x_t, x_{t+1}) - p_t x_t + p_{t+1}x_{t+1} &\geq \\ \delta^t u(x, y) - p_t x + p_{t+1}y \text{ for all } (x, y) &\in \Omega \end{aligned} \tag{4.19}$$

It remains to show that $p_{t+1} \geq 0$. If $x_{t+1} \in [0, b)$, then by choosing $y' \in X$, $y' > x_{t+1}$ in (4.18), we get $p_{t+1} \geq 0$ by the (weak) monotonicity of V on X (which follows from A5). If $x_{t+1} = b$, then by choosing $(x, y) = (x_t, y)$ with $y \in [0, b)$ in (4.19), we get $p_{t+1} \geq 0$ by A5.

We can now provide the proof of Proposition 4.2.

Proof (of Proposition 4.2): Using Lemma 4.1, there is $p_0 \in \Re_+$ satisfying

$$V(x_0) - p_0 x_0 \geq V(x) - p_0 x \text{ for all } x \in X$$

Using Lemma 4.2, there is a sequence $(p_t)_0^\infty$ with $p_t \in \Re_+$ for $t \in \mathbf{N}$ satisfying (4.4) and (4.5)

Using $x = 0$ in (4.5), we get

$$\delta^t[V(x_t) - V(0)] \geq p_t x_t \text{ for } t \in \mathbf{N}$$

Since $|V(x)|$ is bounded on X by $B/(1-\delta)$, and $\delta^t \to 0$ as $t \to \infty$, and $p_t x_t \geq 0$ for $t \in \mathbf{N}$, we have $p_t x_t \to 0$ as $t \to \infty$, establishing (4.6).

5. Sensitivity Analysis

Sensitivity analysis is concerned with studying the nature of the change in the (optimal) policy function, given a change in the parameters of the dynamic optimization problem.

In the present context, the parameters of the dynamic optimization problem (2.1)-(2.3) are (Ω, u, δ) and the initial state, x.

Our treatment of sensitivity analysis will confine itself to the case where the transition possibility set, Ω (and the state space, X) are fixed. We will then study the effect of changes in (u, δ) and the initial state, x, on the nature of the policy function.

We will be concerned with two types of sensitivity questions:

(i) does h change *continuously* with a change in the parameters?

(ii) does h change *monotonically* with a change in the parameters?

We have already seen (in Section 3) that $h(x)$ changes continuously with respect to changes in the initial state x. Thus, in the present section, we will address question (i) above with respect to variations in (u, δ).

It is difficult to address (even formulate) question (ii) above with respect to variations in the utility function, u. Thus, we will address question (ii) with respect to variations in δ and the initial state, x.

5.1 Continuity

The principal result of this subsection is that the policy function is uniformly continuous on the space of all (u, δ), endowed with the metric of uniform convergence. Our exposition is based on Mitra and Sorger (1999).

To proceed, given the space $\mathcal{M}$ of (Ω, u, δ) satisfying A1-A4, define the metric

$$\begin{aligned}\rho((\Omega,u,\delta),(\Omega,u',\delta')) &= \max\{|u(x,y)-u'(x,y)| \\ &: (x,y)\in\Omega\} + |\delta-\delta'|\end{aligned}$$

Proposition 5.1: *For every positive integer n, let $u_n : \Omega \to \Re$ be a utility function and let δ_n be a discount factor, such that $(\Omega, u_n, \delta_n) \in M$. Further, let $u : \Omega \to \Re$ be a utility function and δ be a discount factor with $(\Omega, u, \delta) \in M$, such that*

$$\rho((\Omega,u_n,\delta_n),(\Omega,u,\delta)) \to 0 \quad \text{as } n \to \infty$$

Denote by V_n and h_n [V and h] the value and policy functions for (Ω, u_n, δ_n) [(Ω, u, δ)]. Then V_n converges uniformly to V and h_n converges uniformly to h.

Proof: Since u_n converges uniformly to u, it follows that the sequence $(u_n)_{n=1}^{+\infty}$ is uniformly bounded. Therefore we can find real numbers m and M such that $m \le u(x,y) \le M$ and $m \le u_n(x,y) \le M$ for all $(x,y) \in \Omega$ and all $n \in \{1,2,...\}$. Without loss of generality we may assume that $m = 0$.

Step 1: We start by proving that V_n converges uniformly to V. Let $\varepsilon > 0$ be given. There exists $T \ge 1$ such that

$$\frac{M}{1-\delta}\left(\frac{1+\delta}{2}\right)^{T+1} \le \frac{\varepsilon}{4}.$$

Because of the convergence of the sequences $(\delta_n)_{n=1}^{+\infty}$ and $(u_n)_{n=1}^{+\infty}$ we can find an integer N such that for all $n \ge N$ the following three properties hold:

$$\begin{aligned}\sup\{|u_n(x,y)-u(x,y)| &: (x,y)\in\Omega\} < \varepsilon(1-\delta)/8, \\ \delta_n &\le (1+\delta)/2, \\ \max\{\left|\delta_n^t-\delta^t\right| &: t\in\{1,2,...,T\}\} \le \varepsilon/(4MT).\end{aligned}$$

Consider an arbitrary state $x \in X$ and fix an arbitrary $n \ge N$. Let $(x_t)_{t=0}^{+\infty}$ be an optimal path from initial state x for (Ω, u, δ).

Note that this path is also feasible for the model (Ω, u_n, δ_n). The following chain of inequalities holds because of the above properties:

$$\begin{aligned}
V_n(x) &\geq \sum_{t=0}^{+\infty} \delta_n^t u_n(x_t, x_{t+1}) \\
&\geq \sum_{t=0}^{T} \delta_n^t u_n(x_t, x_{t+1}) \\
&= \sum_{t=0}^{T} (\delta_n^t - \delta^t) u_n(x_t, x_{t+1}) + \sum_{t=0}^{T} \delta^t u_n(x_t, x_{t+1}) \\
&\geq -(\varepsilon/4) + \sum_{t=0}^{T} \delta^t u_n(x_t, x_{t+1}) \\
&\geq -(\varepsilon/4) + \sum_{t=0}^{T} \delta^t u(x_t, x_{t+1}) - (\varepsilon/8) \\
&> \sum_{t=0}^{+\infty} \delta^t u(x_t, x_{t+1}) - \sum_{t=T+1}^{+\infty} \delta^t u(x_t, x_{t+1}) - (\varepsilon/2) \\
&\geq V(x) - (\varepsilon/4) - (\varepsilon/2) \\
&> V(x) - \varepsilon.
\end{aligned}$$

Using an analogous calculation one can also show that $V(x) > V_n(x) - \varepsilon$. Since both $n \geq N$ and $x \in X$ have been chosen arbitrarily, we have shown that V_n converges uniformly to V.

Step 2: Now we prove that h_n converges uniformly to h. To this end first note that h and h_n are continuous functions for all n and that X is compact. If h_n would not converge uniformly to h then, by Royden (1988, p. 162, Exercise 40e) it would be possible to find $x_0 \in X$, $\theta > 0$, and a sequence $(x_n)_{n=1}^{+\infty}$ such that $\lim_{n\to+\infty} x_n = x_0$ and $|h_n(x_n) - h(x_0)| \geq \theta$ for all n. By compactness of X we may assume without loss of generality that $z_0 = \lim_{n\to+\infty} h_n(x_n)$ exists. Since $z_0 \neq h(x_0)$ one can find $\varepsilon > 0$ such that

$$V(x_0) \geq u(x_0, z_0) + \delta V(z_0) + \varepsilon.$$

Because of uniform convergence of u_n to u and V_n to V, convergence of δ_n to δ, as well as continuity of u and V, one can find an intetger N such that for all $n \geq N$ and all $x \in X$ the following properties hold:

$$\begin{aligned}
|V_n(x) - V(x)| &< \varepsilon/8, \\
|V(x_n) - V(x_0)| &< \varepsilon/8, \\
|u(x_n, h_n(x_n)) - u(x_0, z_0)| &< \varepsilon/8, \\
|V(h_n(x_n)) - V(z_0)| &< \varepsilon/8, \\
|\delta_n - \delta| &< \varepsilon(1-\delta)/(8M).
\end{aligned}$$

From these conditions we obtain

$$\begin{aligned}
V_n(x_n) &= [V_n(x_n) - V(x_n)] + [V(x_n) - V(x_0)] + V(x_0) \\
&\geq -(\varepsilon/8) - (\varepsilon/8) + V(x_0) \\
&\geq u(x_0, z_0) + \delta V(z_0) + (3\varepsilon/4) \\
&\geq u(x_n, h_n(x_n)) + \delta V(h_n(x_n)) + (\varepsilon/2) \\
&= u(x_n, h_n(x_n)) + \delta V(h_n(x_n)) + (\varepsilon/2) + (\delta - \delta_n) \\
&\quad V(h_n(x_n)) \\
&\geq u(x_n, h_n(x_n)) + \delta_n V(h_n(x_n)) + (3\varepsilon/8) \\
&\geq u(x_n, h_n(x_n)) + \delta_n V_n(h_n(x_n)) + (\varepsilon/4) \\
&= V_n(x_n) + (\varepsilon/4).
\end{aligned}$$

Clearly, this is a contradiction and the result is proved. Q.E.D.

5.2 Monotonicity with Respect to the Initial State

The concept that is crucial to establishing monotonicity properties of the policy funciton is known as *supermodularity*, and was introduced into the optimization theory literature by Topkis (1978).

Let A be a subset of $\Re^2$ and f a function from A to $\Re$. Then f is *supermodular* on A if whenever (a, b) and (a', b') belong to A with $(a', b') \geq (a, b)$, we have

$$f(a, b) + f(a', b') \geq f(a, b') + f(a', b) \tag{5.1}$$

provided (a, b') and (a', b) belong to A.

If A is a rectangular region, then whenever (a, b) and (a', b') belong to A, we have (a, b') and (a', b) also in A. Further, for such a region if f is continuous on A and C^2 on *int* A, then

$$f_{12}(a, b) \geq 0 \text{ for all } (a, b) \in \mathit{int}\ A \tag{5.2}$$

is equivalent to the condition that f is supermodular on A. [For this result, see Ross (1983), Benhabib-Nishimura (1985).]

Let (Ω, u, δ) be a dynamic optimization model. The principal result on the montonicity of its policy function, h, with respect to the initial condition (x), is that if u is supermodular on Ω, then h is monotone non-decreasing in x. [This result is based on the analysis in Topkis (1978), Benhabib and Nishimura (1985).]

Proposition 5.2: *Let (Ω, u, δ) be a dynamic optimization model, with value function V and policy function h. If u is supermodular on Ω, then h is montone non-decreasing on X.*

Proof: Let x, x' belong to X with $x' > x$. Denote $h(x)$ by z and $h(x')$ by z'. We want to show that $z' \geq z$. Suppose, on the contrary that $z' < z$.

Since $(x, z) \in \Omega$ and $z' \in X$ with $z' \leq z$, we have $(x, z') \in \Omega$. Since $(x, z') \in \Omega$ and $x' \in X$ with $x' \geq x$, we have $(x', z) \in \Omega$. Using the definition of supermodularity, and $(x', z) \geq (x, z')$, we have

$$u(x, z') + u(x', z) \geq u(x, z) + u(x', z') \tag{5.3}$$

Since $z = h(x)$ and $z' = h(x')$, we have

$$V(x) = u(x, z) + \delta V(z) \tag{5.4}$$

and

$$V(x') = u(x', z') + \delta V(z') \tag{5.5}$$

Since $(x, z') \in \Omega$ and $(x', z) \in \Omega$, and $z' \neq z$, we have $z' \neq h(x)$, $z \neq h(x')$, so

$$V(x) > u(x, z') + \delta V(z') \tag{5.6}$$

and

$$V(x') > u(x', z) + \delta V(z) \tag{5.7}$$

Adding (5.4) and (5.5),

$$V(x) + V(x') = u(x, z) + u(x'z') + \delta V(z) + \delta V(z') \tag{5.8}$$

Adding (5.6) and (5.7)

$$V(x) + V(x') > u(x, z') + u(x', z) + \delta V(z') + \delta V(z) \tag{5.9}$$

Using (5.8) and (5.9),

$$u(x, z) + u(x', z') > u(x, z') + u(x', z)$$

which contradicts (5.3) and establishes the result. Q.E.D.

Remark 5.1: It follows from the above result that, in the framework of Proposition 5.2, if (x_t) is an optimal program starting from $x \in X$, then either (i) $x_{t+1} \geq x_t$ for all $t \in \mathbf{N}$ or (ii) $x_{t+1} \leq x_t$ for all $t \in \mathbf{N}$.

5.3 Monotonicity with Respect to the Discount Factor

Let (Ω, u) be given with u supermodular on Ω. Then, for every specification of the discount factor, δ, and initial state x, we can solve the dynamic optimization problem (2.1)-(2.3) and obtain the value $V(\delta, x)$ and policy $h(\delta, x)$.

It turns out that (i) V is supermodular as a function of the two variables (δ, x), and (ii) it follows from this fact that h is monotone non-decreasing in δ.

The second step of this two-step result is fairly straightfoward, using arguments similar to those employed in Section 5.2. The first step is more involved because one cannot use the functional equation of dynamic programming directly to obtain this property of V (since V appears on both sides of the equation). Thus, the method employed is an iterative procedure, in which a sequence of functions, V_n, are shown to be supermodular, and to be converging to the value function, V (Proposition 5.3). In order to make this iterative procedure work, we need to make sure that the property of supermodularity of the maximand is preserved by the maximum value in a maximization problem, a general result due to Topkis (1978). We provide a proof of this result as well (see Lemma 5.1 below) in our special context.

Our analysis in this section is based on Dutta (1987) and Amir, Mirman and Perkins (1991), where similar results are obtained in the more special context of the one-sector model of optimal growth (Example 2.1 of Section 2). Our exposition follows Mitra and Nishimura (1999) closely.

Lemma 5.1: *Let (Ω, u) be given, with u supermodular on Ω. Let $G(\delta, x)$ be a supermodular function on $I \times X$, where $I = (0,1)$, which is continuous on $I \times X$ and non-decreasing and concave on X. Let H be a funciton on $I \times X$ defined by*

$$H(\delta, x) = \max_{z \in \Omega_x} [u(x, z) + \delta G(\delta, z)] \tag{5.10}$$

Then H is supermodular and continuous on $I \times X$, and non-decreasing and concave on X.

Proof: Given the information on G, the maximization problem on the right hand side of (5.1) has a solution. Thus H is well-defined. It is clearly non-decreasing on X.

Using the concavity of u on Ω and G on X, H is concave on X. The continuity of H on $I \times X$ can be established by using arguments similar to those employed in the proof of Proposition 3.1.

It remains to check the supermodularity of H on $I \times X$. Let (δ, x) and (δ', x') belong to $I \times X$ with $(\delta', x') \geq (\delta, x)$. Let $\tilde{z}$ [*resp.*$\tilde{z}'$] be a solution to the maximization problem (5.1) corresponding to (δ', x) [*resp.*(δ, x')].

There are two cases to consider: (i) $\tilde{z}' \geq \tilde{z}$ (ii) $\tilde{z} > \tilde{z}'$.

Case (i) $[\tilde{z}' \geq \tilde{z}]$: From the definitions of $\tilde{z}$ and $\tilde{z}'$ it follows that

$$H(\delta, x') = u(x', \tilde{z}') + \delta G(\delta, \tilde{z}') \tag{5.11}$$

and

$$H(\delta', x) = u(x, \widetilde{z}) + \delta' G(\delta', \widetilde{z}) \tag{5.12}$$

Also, considering the maximization problem (5.10) for (δ, x) and (δ', x') yield:

$$H(\delta, x) \geq u(x, \widetilde{z}) + \delta G(\delta, \widetilde{z}) \tag{5.13}$$

and

$$H(\delta', x') \geq u(x', \widetilde{z}') + \delta' G(\delta', \widetilde{z}') \tag{5.14}$$

Since $(\delta', \widetilde{z}') \geq (\delta, \widetilde{z})$ and G is supermodular, we get

$$G(\delta', \widetilde{z}') + G(\delta, \widetilde{z}) \geq G(\delta', \widetilde{z}) + G(\delta, \widetilde{z}')$$

which yields

$$G(\delta', \widetilde{z}') - G(\delta', \widetilde{z}) \geq G(\delta, \widetilde{z}') - G(\delta, \widetilde{z}) \geq 0 \tag{5.15}$$

using the fact that G is non-decreasing on X. Thus, we get

$$\begin{aligned} \delta'[G(\delta', \widetilde{z}') - G(\delta', \widetilde{z})] &\geq \delta'[G(\delta, \widetilde{z}') - G(\delta, \widetilde{z})] \\ &\geq \delta[G(\delta, \widetilde{z}') - G(\delta, \widetilde{z})] \end{aligned}$$

which in turn yields

$$\delta' G(\delta', \widetilde{z}') + \delta G(\delta, \widetilde{z}) \geq \delta' G(\delta', \widetilde{z}) + \delta G(\delta, \widetilde{z}') \tag{5.16}$$

Adding (5.11) and (5.12), and (5.13) and (5.14), and using (5.16), we get

$$H(\delta, x) + H(\delta', x') \geq H(\delta, x') + H(\delta', x)$$

which shows that H is supermodular.

Case (ii) $[\widetilde{z} > \widetilde{z}']$: Note that $(x, \widetilde{z}) \in \Omega$ and $\widetilde{z} > \widetilde{z}'$, so that $(x, \widetilde{z}') \in \Omega$. Also, $(x, \widetilde{z}) \in \Omega$ and $x' \geq x$, so that $(x', \widetilde{z}) \in \Omega$. Thus, considering the maximization problem (5.10) for (δ, x) and (δ', x') yield:

$$H(\delta, x) \geq u(x, \tilde{z}') + \delta G(\delta, \tilde{z}') \tag{5.17}$$

$$H(\delta', x') \geq u(x', \tilde{z}) + \delta' G(\delta', \tilde{z}) \tag{5.18}$$

Also, note that (5.11) and (5.12) continue to be valid. Adding (5.11) and (5.12), and (5.17) and (5.18), and using the supermodularity of u, we get

$$H(\delta, x) + H(\delta', x') \geq H(\delta, x') + H(\delta', x)$$

which establishes the supermodularity of H. Q.E.D.

Proposition 5.3: *Let (Ω, u) be given with u supermodular on Ω. Then (i) $V(\delta, x)$ is supermodular on $I \times X$, and (ii) $h(\delta, x)$ is monotone non-decreasing on I.*

Proof: Normalize $u(0,0) = 0$, and define:

$$V_0(\delta, x) = u(x, 0) \tag{5.19}$$

$$V_{n+1}(\delta, x) = \max_{z \in \Omega_x} [u(x, z) + \delta V_n(\delta, z)] \tag{5.20}$$

Note that $V_0(\delta, x)$ is continuous on $I \times X$ and non-decreasing and concave on X. Further, it is (trivially) supermodular. Thus, using Lemma 5.1, (5.20) defines a sequence $\{V_n(\delta, x)\}$ of functions on $I \times X$, such that V_n is continuous and supermodular on $I \times X$, and non-decreasing and concave on X.

Since u is bounded on X, we can find $B > 0$ such that $|u(x, z)| \leq B$ for all $(x, z) \in \Omega$. Then $|V_n(\delta, x)| \leq B/(1 - \delta)$ for each $\delta \in I$, as can be easily verified by induction.

Next, we show that $V_{n+1}(\delta, x) \geq V_n(\delta, x)$ for $n = 0, 1, 2, \ldots$. Notice that $V_1(\delta, x) \geq u(x, 0) + \delta V_0(\delta, 0) = u(x, 0) + \delta u(0, 0) =$

$u(x,0) = V_0(\delta, x)$. Assume now that $V_{n+1}(\delta, x) \geq V_n(\delta, x)$ for $n = 0, ..., N$, where $N \geq 0$. We will now show that $V_{N+2}(\delta, x) \geq V_{N+1}(\delta, x)$. Let $\overline{z}$ be the solution to the problem:

$$\max_{z \in \Omega_x}[u(x,z) + \delta V_N(\delta, z)]$$

Then $V_{N+2}(\delta, x) = \max_{z \in \Omega_x}[u(x,z) + \delta V_{N+1}(\delta, z)]$
$\geq u(x, \overline{z}) + \delta V_{N+1}(\delta, \overline{z}) \geq u(x, \overline{z}) + \delta V_N(\delta, \overline{z}) = V_{N+1}(\delta, x)$, by definition of $\overline{z}$.

Fix any δ in I, and consider the sequence of functions $\{V_n(\delta, \bullet)\}_{n=0}^{\infty}$. Define $\overline{V}(\delta, x) = \lim_{n \to \infty} V_n(\delta, x)$ for each $x \in X$, noting that the limit is well defined. Then (using Theorem 7.13, p. 150 of Rudin (1976)) $V_n(\delta, \bullet)$ in fact converges uniformly to $\overline{V}(\delta, \bullet)$, and so $\overline{V}(\delta, \bullet)$ is continuous on X. Also, $\overline{V}(\delta, \bullet)$ is concave on X.

Fixing δ in I, denote by z_n the solution to the maximization problem on the right hand side of (5.20). Then, there is a subsequence of $\{z_n\}$ which converges to some $\overline{z}$. Using the fact that

$$V_{n+1}(\delta, x) = u(x, z_n) + \delta V_n(\delta, z_n)$$

we obtain

$$\overline{V}(\delta, x) \leq u(x, \overline{z}) + \delta \overline{V}(\delta, \overline{z}) \tag{5.21}$$

Also, for every $z \in \Omega_x$, we have $V_{n+1}(\delta, x) \geq u(x, z) + \delta V_n(\delta, z)$, so that

$$\overline{V}(\delta, x) \geq u(x, z) + \delta \overline{V}(\delta, z) \tag{5.22}$$

Using (5.21) and (5.22), we get

$$\overline{V}(\delta, x) = \max_{z \in \Omega_x}[u(x,z) + \delta \overline{V}(\delta, z)]$$

Since $\overline{V}(\delta, \bullet)$ is continuous on X, we must have $\overline{V}(\delta, \bullet) = V(\delta, \bullet)$ by Proposition 3.1.

Since V_n is supermodular on $I \times X$, and $V_n(\delta, x) \to V(\delta, x)$ as $n \to \infty$, for all (δ, x) in $I \times X$, V is supermodular on $I \times X$. This establishes (i).

To establish (ii), let $x \in X$, and $\delta,\ \delta' \in I$ with $\delta' > \delta$. Denote $h(\delta, x)$ by z and $h(\delta', x)$ by z'. We want to show that $z' \geq z$. Suppose, on the contrary, that $z' < z$.

Since $z = h(\delta, x)$ and $z' = h(\delta', x)$, we get

$$V(\delta, x) = u(x, z) + \delta V(\delta, z) \tag{5.23}$$

and

$$V(\delta', x) = u(x, z') + \delta' V(\delta', z') \tag{5.24}$$

Also, we clearly have (since $z \neq z'$),

$$V(\delta, x) > u(x, z') + \delta V(\delta, z') \tag{5.25}$$

and

$$V(\delta', x) > u(x, z) + \delta' V(\delta', z) \tag{5.26}$$

Using (5.23)-(5.26), we have

$$\delta V(\delta, z') + \delta' V(\delta', z) \leq \delta V(\delta, z) + \delta' V(\delta', z')$$

which yields

$$\delta'[V(\delta', z) - V(\delta', z')] \leq \delta[V(\delta, z) - V(\delta, z')] \tag{5.27}$$

Since $z > z'$, $[V(\delta', z) - V(\delta', z')] \geq 0$, so that, using $\delta' > \delta$, we get

$$\delta'[V(\delta', z) - V(\delta', z')] \geq \delta[V(\delta', z) - V(\delta', z')] \tag{5.28}$$

Using (5.27) and (5.28) we get

$$V(\delta', z) - V(\delta', z') \leq V(\delta, z) - V(\delta, z')$$

which yields

$$V(\delta', z) + V(\delta, z') \leq V(\delta', z') + V(\delta, z) \tag{5.29}$$

But since $\delta' > \delta$ and $z > z'$, (5.29) violates the supermodularity of V. Q.E.D.

6. Existence of a Stationary Optimal Stock

The concept of a (non-trivial) stationary optimal stock plays a central role in the theory of optimal intertemporal allocation. Its primary significance derives from the fact that it is the rest point of the dynamical system (X, h), which is a global attractor for optimal programs from all positive initial stocks (usually referred to as the turnpike property) when the discount factor is close to unity.

We will discuss this stability property in Section 9 below. In the present section, we confine ourselves to discussing the existence of a stationary optimal stock. Our exposition is primarily based on Flynn (1980), McKenzie (1982) and Khan and Mitra (1986).

6.1 Existence of a Discounted Golden-Rule Stock

A *discounted golden-rule stock* k is an element of X such that
(i) $(k, k) \in \Omega$
(ii) $u(k, k) \geq u(x, y)$ for all $(x, y) \in \Omega$ such that $\delta y - x \geq \delta k - k$
(iii) $u(k, k) > u(0, 0)$

We establish the existence of a discounted golden-rule stock under a "δ-normality" condition. We say that (Ω, u, δ) is *δ-normal* if there exists $(\overline{x}, \overline{y}) \in \Omega$ such that $\delta\overline{y} \geq \overline{x}$ and $u(\overline{x}, \overline{y}) > u(0, 0)$.

Proposition 6.1: *If (Ω, u, δ) satisfies A1-A3 and A5, and is δ-normal, then there exists a discounted golden-rule stock.*

Proof: For z in X, define $\phi(z) = \{(x, y) \in \Omega : \delta y - x \geq \delta z - z\}$ and $\psi(z) = \{(x, y) \in \Omega : u(x, y) \geq u(x', y')$ for all $(x', y') \in \phi(z)\}$.

Note that for each $z \in X$, the set $\phi(z)$ is non-empty (since $(0, 0) \in \Omega$), compact (since Ω is closed) and convex (since Ω is convex). Thus, for each $z \in X$, the set $\psi(z)$ is non-empty (since u is upper semicontinuous on Ω) and convex (since u is concave).

Next, we show the upper hemicontinuity of ψ. Let z^* be an arbitrary point of X. Consider a sequence $z^n \in X$, with $z^n \to z^*$ as $n \to \infty$. Let $(x^n, y^n) \in \psi(z^n)$, and $(x^n, y^n) \to (\widehat{x}, \widehat{y})$. We want to show that $(\widehat{x}, \widehat{y}) \in \psi(z^*)$. Since Ω is closed, $(\widehat{x}, \widehat{y}) \in \phi(z^*)$. Suppose $(\widehat{x}, \widehat{y})$ is not in $\psi(z^*)$. Then there is some $(x^*, y^*) \in \psi(z^*)$ and $\varepsilon > 0$ such that $u(x^*, y^*) \geq u(\widehat{x}, \widehat{y}) + \varepsilon$.

Now, since u is an upper semicontinuous function, $\limsup_{n\to\infty} u(x^n, y^n) \le u(\widehat{x}, \widehat{y})$. Thus, there is N_1 such that for $n \ge N_1$, $u(x^n, y^n) \le u(\widehat{x}, \widehat{y}) + \varepsilon/3$. Consequently, for $n \ge N_1$,

$$u(x^*, y^*) \ge u(x^n, y^n) + 2\varepsilon/3 \tag{6.1}$$

Choose $0 < \lambda < 1$ such that $(1-\lambda)[u(0,0) - u(x^*, y^*)] \ge -\varepsilon/3$. We claim that there is an N_2 such that for $n \ge N_2$, $(\lambda x^*, \lambda y^*) \in \phi(z^n)$. To see this, observe that $(0,0) \in \Omega$ and convexity of Ω imply that $(\lambda x^*, \lambda y^*) \in \phi(\lambda z^*)$. Since $z^n \to z^*$, there is N_2 such that for $n > N_2$, $z^n \ge \lambda z^*$. Thus $\delta\lambda y^* - \lambda x^* \ge (\delta - 1)\lambda z^* \ge (\delta - 1) z^n$, establishing our claim.

Since $(x^n, y^n) \in \psi(z^n)$, we have for $n \ge N_2$,

$$\begin{aligned} u(x^n, y^n) &\ge u(\lambda x^*, \lambda y^*) \ge \lambda u(x^*, y^*) + (1-\lambda) u(0,0) \\ &= u(x^*, y^*) + (1-\lambda)[u(0,0) - u(x^*, y^*)] \\ &\ge u(x^*, y^*) - \varepsilon/3. \end{aligned}$$

Using this in (6.1) for $n \ge Max(N_1, N_2)$,

$$u(x^*, y^*) \ge u(x^n, y^n) + 2\varepsilon/3 \ge u(x^*, y^*) + \varepsilon/3,$$

which leads to a contradiction and completes the demonstration that ψ is upper hemicontinuous.

Define for $z \in X$, $Q(z) = \{x \in X : (x, y) \in \psi(z)\}$. We will show that this correspondence Q satisfies all the requirements of Kakutani's fixed-point theorem.

Clearly, Q is a non-empty, convex-valued correspondence from X to the subsets of X. We check that Q is upper hemicontinuous. To see this, take an arbitrary $z^* \in X$. Let $z^n \in X$, with $z^n \to z^*$ as $n \to \infty$. Let $x^n \in Q(z^n)$, and $x^n \to \widehat{x}$ as $n \to \infty$. We have to show that $\widehat{x} \in Q(z^*)$. Since $x^n \in Q(z^n)$, there is y^n such that $(x^n, y^n) \in \psi(z^n)$. This means $(x^n, y^n) \in \phi(z^n)$, and we can pick a subsequence $(x^{n'}, y^{n'})$ tending to $(\widehat{x}, \widehat{y}) \in \phi(z^*)$. Since ψ is upper hemicontinuous, $(\widehat{x}, \widehat{y}) \in \psi(z^*)$ and so $\widehat{x} \in Q(z^*)$.

Thus, all the conditions of Kakutani's fixed point theorem are fulfilled, and there exists $x^o \in Q(x^o)$. This means there is some y^o such that $(x^o, y^o) \in \psi(x^o)$. That is,

$$u(x^o, y^o) \ge u(x, y) \text{ for all } (x, y) \in \phi(x^o).$$

But $(x^o, y^o) \in \phi(x^o)$ implies $x^o \leq y^o$, and we obtain from A5 that $(x^o, x^o) \in \Omega$ and $u(x^o, x^o) \geq u(x^o, y^o) \geq u(x, y)$ for all $(x, y) \in \Omega$ with $\delta y - x \geq \delta x^o - x^o$. Given δ-normality, there is $(x', y') \in \phi(x^o)$ such that $u(x', y') > u(0, 0)$. Thus $u(x^o, x^o) > u(0, 0)$, and hence x^o is a discounted golden-rule stock.

6.2 Existence of a Non-Trivial Stationary Optimal Stock

An optimal program $(x(t))_0^\infty$ from k is a *stationary optimal program* if $x(t) = k$ for $t \in \mathbf{N}$. A *stationary optimal stock* k is an element of X such that $(k)_0^\infty$ is a stationary optimal program. It is called *non-trivial* of $u(k, k) > u(0, 0)$.

Using Proposition 6.1, we can establish the existence of a non-trivial stationary optimal stock.

Proposition 6.2: *If (Ω, u, δ) satisfies A1-A3 and A5, and is δ-normal, then there exists a non-trivial stationary optimal stock.*

Proof: By Proposition 6.1, there is a discounted golden-rule stock, k. Now, let $(x_t)_0^\infty$ be any program from k. We will show that

$$\sum_{t=0}^{\infty} \delta^t u(k, k) \geq \sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}) \tag{6.2}$$

Let $X_T = \sum_{t=0}^{T-1} (1-\delta)\delta^t x_t/(1-\delta^T)$ and $Y_T = \sum_{t=0}^{T-1} (1-\delta)\delta^t x_{t+1}/(1-\delta^T)$ for $T = 1, 2, 3, \ldots$. Given convexity of Ω, we have (X_T, Y_T) in Ω for $T = 1, 2, 3, \ldots$. Since $x_t \in X$ for all $t \in \mathbf{N}$, the pair $(\overline{x}, \overline{y}) \equiv \lim_{T\to\infty} (X_T, Y_T)$ is well-defined and in Ω.

Now, by concavity of u, and the fact that $0 < \delta < 1$, Jensen's inequality yields $u(\overline{x}, \overline{y}) \geq \sum_{t=0}^{\infty} (1 - \delta)\delta^t u(x_t, x_{t+1})$. But $(\overline{x} - \delta\overline{y}) = (1 - \delta) \left[\sum_{t=0}^{\infty} \delta^t x_t - \sum_{t=0}^{\infty} \delta^{t+1} x_{t+1} \right] = (1 - \delta)k$. Since k is a discounted golden-rul stock, $u(k,k) \geq u(\overline{x},\overline{y})$, which in turns implies that

$$\sum_{t=0}^{\infty} \delta^t u(k,k) \geq \sum_{t=0}^{\infty} \delta^t u(\overline{x},\overline{y}) = u(\overline{x},\overline{y})/(1-\delta) \geq \sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1})$$

which establishes (6.2). Thus, k is a stationary optimal stock. Since k is a discounted golden-rule stock, we have $u(k,k) > u(0,0)$, and so k is a non-trivial stationary optimal stock.

6.3 The Existence of a Modified Golden-Rule

A *modified golden-rule* is a pair (k,p) such that
(i) $(k,k) \in \Omega$, $p \in \Re_+$
(ii) $u(x,y) + \delta py - px \leq u(k,k) + \delta pk - pk$ for all $(x,y) \in \Omega$
Given Proposition 6.1, we can establish the existence of a modified golden-rule by providing a price support to the discounted golden-rule stock.

Proposition 6.3: *If (Ω, u, δ) satisfies A1-A3 and A5, and is δ-normal, then there exists a modified golden-rule.*

Proof: By Proposition 6.1, there is a discounted golden-rule stock, k.

Define two sets A and B as follows:
$A = \{(a,b) \in \Re^2 : u(x,y) - u(k,k) \geq a,\ (\delta y - x) - (\delta k - k) \geq b$ for some $(x,y) \in \Omega\}$
$B = \{(a,b) \in \Re^2 : a > 0 \text{ and } b > 0\}$

Then A and B are non-empty, convex sets. By the definition of a discounted golden-rule stock, A and B are disjoint. Then, by a separation theorem (see Nikaido (1968), Theorem 3.5, p. 35), there is $(\mu, v) \in \Re^2_+$, $(\mu, v) \neq 0$ such that

$$\mu a + vb \leq 0 \text{ for all } (a,b) \in A$$

We claim that $\mu \neq 0$. For if $\mu = 0$, then we have $vb \leq 0$ for all $(a,b) \in A$. Define $a = u(0,0) - u(k,k)$ and $b = (k - \delta k)$. Since $(0,0) \in \Omega$, $(a,b) \in A$, and so $v(k - \delta k) \leq 0$. Since $u(k,k) > u(0,0)$, we have $k > 0$ and so $(k - \delta k) > 0$. Thus, $v \leq 0$, which means $v = 0$, contradicting the fact that $(\mu, v) \neq 0$. Thus, $\mu > 0$, and so

$$a + (v/\mu)b \leq 0 \text{ for all } (a, b) \in A.$$

Defining $p = (v/\mu)$, we have

$$u(x, y) + \delta py - px \leq u(k, k) + \delta pk - pk \text{ for all } (x, y) \in \Omega$$

Thus, (k, p) is a modified golden-rule.

7. Smooth Preferences

So far, we have developed the theory of dynamic optimization without assumptions on the smoothness of the reduced form utility function.

However, significant progress in the theory can be made (beyond the material covered in Sections 2-6) if one assumes that the utility function is smooth in the interior of the transition possibility set.

We will concentrate on two sets of ideas that emerge from this theory. First, if the utility function, u, is C^1 in the interior of Ω, then an interior optimal program satisfies the Ramsey-Euler equations. Further, the value function is continuously differentiable in the interior of the state space, and the derivative of the value function is the derivative of the utility function (with respect to its first argument) when evaluated at the optimal point (by the envelope theorem).

Going back to the material on duality theory covered in Section 5, we now see that the shadow prices supporting an optimal program are given exactly by the derivatives of the value function (at the appropriate optimal states). This leads to a restatement of the characterization of optimal programs in terms of the Ramsey-Euler equations and a transversality condition that is quite useful. In particular, any interior stationary (and, in fact, any periodic) optimal program can be characterized solely in terms of the Ramsey-Euler equations.

The second set of ideas is developed under the hypotheses that the utility function, u, is C^2 in the interior of Ω, and optimal programs are in the interior of Ω (when the initial state is in the interior of X). Now, the Ramsey-Euler equations involve continuously differentiable functions which can be linearized around an interior stationary optimal stock to study the local stability of optimal programs in terms of the characteristic roots associated with the Ramsey-Euler equation at the stationary optimal stock.

The interesting feature that emerges is that the dynamic behavior of optimal programs in governed by the smaller (in absolute value) of the characteristic roots associated with the Ramsey-Euler equation. This feature is made precise in the result that (i) the policy function is differentiable at the stationary optimal stock, and (ii) the derivative of the policy function at the stationary optimal stock is equal to the smaller (in absolute value) of the characteristic roots associated with the Ramsey-Euler equation at the stationary optimal stock.

7.1 Ramsey-Euler Equations and the Differentiability of the Value Function

Ramsey-Euler equations can be derived as first-order necessary conditions of the optimization problem (2.1)-(2.3). Thus, if (x_t^*) solves the optimization problem (2.1)-(2.3), then for each $t \in \mathbf{N}$, x_{t+1}^* must solve the problem:

$$(Q)\begin{cases} \text{Max } u(x_t^*, y) + \delta u(y, x_{t+2}^*) \\ \text{subject to } y \in \Omega_{x_t^*}, x_{t+2}^* \in \Omega_y \end{cases}$$

Introduce now the following smoothness assumption on u:

A6: u is continuously differentiable in the interior of Ω.

If the optimal program is in the interior of Ω for all t (that is, if $(x_t^*, x_{t+1}^*) \in \text{int } \Omega$ for all $t \in \mathbf{N}$), then the first-order conditions for (Q) are:

$$u_2(x_t^*, x_{t+1}^*) + \delta u_1(x_{t+1}^*, x_{t+2}^*) = 0 \text{ for } t \in \mathbf{N} \tag{7.1}$$

which are the Ramsey-Euler equations.

We summarize this finding in the following result.

Proposition 7.1: *Suppose the policy function, h, satisfies $(x, h(x)) \in$ int Ω for all $x \in$ int X, then*

$$u_2(x, h(x)) + \delta u_1(h(x), h^2(x)) = 0$$

for all $x \in$ int X.

If an optimal program (x_t) is interior (that is, $(x_t, x_{t+1}) \in \text{int}\ \Omega$ for all $t \in \mathbf{N}$), then the shadow prices (p_t) obtained in Proposition 4.2 must satisfy (using (4.4))

$$\delta^t u_1(x_t, x_{t+1}) = p_t$$

If V were differentiable, then we would also have from (4.5)

$$\delta^t V'(x_t) = p_t$$

Thus, when V is differentiable, we should have (for $x \in \text{int}\ X$),

$$V'(x) = u_1(x, h(x))$$

We now show that (under A7), V is continuously differentiable in int X, and the above formula, relating the derivatives of the value and utility functions, holds. This result is due to Benveniste and Scheinkman (1979).

Proposition 7.2: *Suppose the policy function, h, satisfies $(x, h(x)) \in$ int Ω for all $x \in$ int X. Then V is continuously differentiable on int X, and*

$$V'(x) = u_1(x, h(x)) \tag{7.2}$$

for all $x \in$ int X.

Proof: Pick any $x^0 \in \text{int}\ X$, and note that by assumption $(x^0, h(x^0)) \in \text{int}\ \Omega$. Then there is a neighborhood N of x^0 such that $(x, h(x^0)) \in \text{int}\ \Omega$ for all $x \in N$. Define W on N by

$$W(x) = u(x, h(x^0)) + \delta V(h(x^0)) \tag{7.3}$$

Then W is concave and continuously differentiable on N. Further, $W(x^0) = V(x^0)$, and for all $x \in N$, $W(x) = u(x, h(x^0)) + \delta V(h(x^0)) \leq V(x)$. Let p be any subgradient of V at x^0. Then

$$p(x - x^0) \geq V(x) - V(x^0) \geq W(x) - W(x^0)$$

Since W is concave and continuously differentiable on N, $p = W'(x^0)$. Thus, V has a unique subgradient at x^0, so it must be differentiable at x^0, with $p = V'(x^0)$. Since $x^0 \in \text{int } X$ was arbitrary, V is differentiable on int X. Since V is concave, V is in fact, continuously differentiable on int X.

Using (7.3) to evaluate the derivative of W at x^0, we get

$$W'(x^0) = u_1(x^0, h(x^0)) \tag{7.4}$$

Since $W'(x^0) = p = V'(x^0)$, we get

$$V'(x^0) = u_1(x^0, h(x^0)) \tag{7.5}$$

Since $x^0 \in \text{int } X$ was arbitrary, this establishes (7.2). Q.E.D.

Given the above results, it is now possible to restate the results on the price characterization of optimal programs somewhat differently as follows.

Corollary 7.1: *If $(x_t)_0^\infty$ is a program from $x \in X$, with $(x_t, x_{t+1}) \in$ int Ω for $t \in \mathbf{N}$, and the following conditions hold:*

$$\text{(i) } u_2(x_t, x_{t+1}) + \delta u_1(x_{t+1}, x_{t+2}) = 0 \tag{7.6}$$

$$\text{(ii) } \lim_{t \to \infty} \delta^t u_1(x_t, x_{t+1}) x_t = 0 \tag{7.7}$$

then $(x_t)_0^\infty$ is an optimal program from x.

Proof: Define $p_t = \delta^t u_1(x_t, x_{t+1})$ for $t \in \mathbf{N}$. Note that $p_t \in \Re_+$ for $t \in \mathbf{N}$ by A5. Then for all $(x, z) \in \Omega$ and $t \in \mathbf{N}$,

$$\begin{aligned} u(x, z) - u(x_t, x_{t+1}) \quad \leq \quad & u_1(x_t, x_{t+1})(x - x_t) \\ & + u_2(x_t, x_{t+1})(z - x_{t+1}) \end{aligned}$$

Multiplying through by δ^t, and using (7.6) and the definition of p_t,

$$\delta^t[u(x,z) - u(x_t, x_{t+1})] \leq p_t(x - x_t) - p_{t+1}(z - x_{t+1}) \qquad (7.8)$$

Transposing terms in (7.8) yields (4.1). Also, (7.7) and the definition of p_t yields (4.2). Now the result follows from Proposition 4.1.

Remark 7.1: (i) It is worth noting that unlike Proposition 4.1, the sufficiency result of Corollary 7.1 does exploit the convex structure of the dynamic optimization model.

(ii) It follows from Corollary 7.1 that if x is a state satisfying $(x,x) \in \text{int } \Omega$, and $u_2(x,x) + \delta u_1(x,x) = 0$,
then X is a stationary optimal stock, since the transversality condition (7.7) is automatically satisfied.

(iii) Generalizing the result in remark (ii) above, if $(x_0, x_1, ..., x_N)$ (with $N \geq 1$) is a vector such that $(x_t, x_{t+1}) \in \text{int } \Omega$, and
$u_2(x_t, x_{t+1}) + \delta u_1(x_{t+1}, x_{t+2}) = 0$ for $t = 0, ..., N-2$ (if $N \geq 2$),
$u_2(x_{N-1}, x_N) + \delta u_1(x_N, x_0) = 0$, then $(x_0, x_1, ..., x_N)$ is a periodic optimal program.

A converse of Corollary 7.1 can be established, using Proposition 4.2.

Corollary 7.2: *Suppose $(x_t)_0^\infty$ is an optimal program from $x \in X$, such that $(x_t, x_{t+1}) \in$ int Ω for $t \in \mathbf{N}$. Then the following conditions hold:*

(i) $u_2(x_t, x_{t+1}) + \delta u_1(x_{t+1}, x_{t+2}) = 0$ *for* $t \in \mathbf{N}$

(ii) $\lim_{t\to\infty} \delta^t u_1(x_t, x_{t+1}) x_t = 0$

(iii) $V'(x_t) = u_1(x_t, x_{t+1})$ *for* $t \in \mathbf{N}$

Proof: Using (4.4) and the fact that (x_t, x_{t+1}) is in int Ω for $t \in \mathbf{N}$, we have

$$\delta^t u_1(x_t, x_{t+1}) = p_t \qquad (7.9)$$

$$\delta^t u_2(x_t, x_{t+1}) = -p_{t+1} \tag{7.10}$$

This yields (i). Using (4.6), (ii) follows from (7.9).

By the proof of Proposition 7.2, V is differentiable at x_t for each $t \in \mathbf{N}$. So, using (4.5),

$$\delta^t V'(x_t) = p_t \tag{7.11}$$

Thus, (iii) follows from (7.9) and (7.11).

7.2 Characteristic Roots of the Ramsey-Euler Equation and Differentiability of the Policy Function

The study of the differentiability of the policy function is one of the hard topics of dynamic optimization theory. For the present exposition, we will confine our attention to the differentiability of the policy function at a stationary optimal stock, which turns out to be the most useful aspect of this theory for applications.

The differentiability property of the policy function at an *SOS* can be used to establish that the derivative of the policy function equals the smaller (in absolute value) characteristic root associated with the Ramsey-Euler equation, at the *SOS*. (See Santos (1991) for the most general result).

We will approach the topic in reverse order. That is, we will establish a relation betwenn the Dini derivatives of the policy function and the characteristic roots associated with the Ramsey-Euler equation at the *SOS*. This will then be used to establish simultaneously the differentiability of the policy function at the *SOS*, and the equality of the derivative of the policy function to the smaller (in absolute value) characteristic root associated with the Ramsey-Euler equation. This approach is based on Mitra and Nishimura (1999).

We proceed with our analysis under the following condition:

C1: There is a stationary optimal stock, $k_\delta \in X$, such that $(k_\delta, k_\delta) \in \text{int } \Omega$, $u_{11}(k_\delta, k_\delta) < 0$, $u_{22}(k_\delta, k_\delta) < 0$ and $0 < (1 + \delta)|u_{12}(k_\delta, k_\delta)| < [-u_{22}(k_\delta, k_\delta)] + \delta[-u_{11}(k_\delta, k_\delta)]$.

We will carry out our analysis for the case $u_{12}(k_\delta, k_\delta) > 0$, noting that the analysis for the case $u_{12}(k_\delta, k_\delta) < 0$ is analogous (and is worked out explicitly in Mitra and Nishimura (1999)).

We write k for k_δ to simplify the notation. Since $(k, k) \in$ int Ω, we can find an open interval $N_1(k)$ containing k, such that $N_1(k)^2$ is in the interior of Ω, and for all (x, z) in $N_1(k)^2$, we have $u_{12}(x, z) > 0$. Now, we can find an open interval $N(k)$, containng k, with $N(k) \subset N_1(k)$, such that for all $x \in N(k)$, $h(x)$ and $h^2(x)$ belong to $N_1(k)$.

Then, for all x, x' in $N(k)$, with $x' > x$, we must have $h(x') > h(x)$. To see this, note that we have

$$u_2(x', h(x')) = -\delta V'(h(x'))$$

$$u_2(x, h(x)) = -\delta V'(h(x))$$

Thus, by the Mean-Value Theorem,

$$\begin{aligned} & \widetilde{u}_{21}(x' - x) + \widetilde{u}_{22}(h(x') - h(x)) \\ = \; & \delta[V'(h(x)) - V'(h(x'))] \end{aligned}$$

where $\widetilde{u}_{21}$ and $\widetilde{u}_{22}$ are evaluated at a point (y, z) in between $(x, h(x))$ and $(x'(h(x'))$, as given by the Mean-Value Theorem. If $h(x') \leq h(x)$, then $V'(h(x')) \geq V'(h(x))$, so that

$$\widetilde{u}_{21}(x' - x) + (-\widetilde{u}_{22})(h(x) - h(x')) \leq 0$$

But since $\widetilde{u}_{21} > 0$, $(-\widetilde{u}_{22}) > 0$, we have a contradiction.

We break up our analysis into several steps.

Step 1: We can write the characteristic equation [associated with the Ramsey-Euler equation at $x = k_\delta$] as follows:

$$u_{12}(k, k) + [u_{22}(k, k) + \delta u_{11}(k, k)]\lambda + \delta u_{12}(k, k)\lambda^2 = 0 \quad (7.12)$$

Denoting the roots of the characteristic equation by λ_1 and λ_2 (where λ_1 has the least absolute value), we can infer that k is "saddle-point stable"; that is

$$0 < \lambda_1 < 1 < \lambda_2 \quad (7.13)$$

To see this, observe first that by using C1, λ_1 and λ_2 are real, and

$$\lambda_1\lambda_2 = (1/\delta); (\lambda_1 + \lambda_2) = -\{[u_{22}(k,k) + \delta u_{11}(k,k)]/\delta u_{12}(k,k)\} \tag{7.14}$$

Thus, λ_1 and λ_2 are both positive, with $(\lambda_2)^2 \geq \lambda_1\lambda_2 = (1/\delta)$, so that $\lambda_2 > 1$.

Using (7.14) and C1, we obtain $(1-\lambda_1)(1-\lambda_2) = 1 - \lambda_1 - \lambda_2 + \lambda_1\lambda_2 = [1/\delta u_{12}(k,k)][(1+\delta)u_{12}(k,k) + u_{22}(k,k) + \delta u_{11}(k,k)] < 0$. Since $\lambda_2 > 1$, we must have $\lambda_1 < 1$. This verifies (7.13).

Step 2: If $x \in N(k)$, $x \neq \widehat{k}$, then

$$\widetilde{u}_{21} + [(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11})]\{[h(x) - h(k)]/(k-x)\}$$

$$+\delta\overline{u}_{12}\{[h^2(k) - h^2(x)]/(k-x)\} = 0 \tag{7.15}$$

where $\widetilde{u}_{21}$ and $\widetilde{u}_{22}$ are evaluated at an appropriate convex combination of $(x, h(x))$ and (k,k), and $\overline{u}_{11}$ and $\overline{u}_{12}$ are evaluated at an appropriate convex combination of $(h(x), h^2(x))$ and (k,k) as given by the Mean-Value Theorem.

To see this, write the Ramsey-Euler equations:

$$\begin{aligned} u_2(k,k) + \delta u_1(k,k) &= 0 \\ u_2(x, h(x)) + \delta u_1(h(x), h^2(x)) &= 0 \end{aligned}$$

Use the Mean-Value Theorem to get

$$\widetilde{u}_{21}(k-x) + (-\widetilde{u}_{22})(h(x)-k) + \delta(-\overline{u}_{11})(h(x)-k)$$

$$+\delta\overline{u}_{12}[k - h^2(x)] = 0$$

Dividing by $(k-x) \neq 0$, we obtain (7.15).

Step 3: For $x \in N(k), n \neq k$, define

$$m = \min\left[\liminf_{x\to k+}\left|\frac{h(x)-h(k)}{x-k}\right|, \liminf_{x\to k-}\left|\frac{h(x)-h(k)}{x-k}\right|\right]$$

$$M = \max\left[\limsup_{x\to k^+}\left|\frac{h(x)-h(k)}{x-k}\right|, \limsup_{x\to k^-}\left|\frac{h(x)-h(k)}{x-k}\right|\right]$$

We claim that

(i) M is either λ_1 or λ_2; (ii)m is either λ_1 or λ_2.

We will establish only (i), since the proof of (ii) is similar. Let us define:

$$\begin{aligned} g(\lambda) &= u_{21}(k,k) + u_{22}(k,k)\lambda + \delta u_{11}(k,k)\lambda \\ &\quad + \delta u_{12}(k,k)\lambda^2 \end{aligned}$$

Claim 1: $g(M) > 0$ is not possible. For if $g(M) > 0$, then

$$u_{21}(k,k) + [u_{22}(k,k) + \delta u_{11}(k,k)]M + \delta u_{12}(k,k)M^2 > 0$$

We can choose $\varepsilon > 0$ such that

$$\begin{aligned} u_{21}(k,k) &> [(-u_{22}(k,k)) + \delta(-u_{11}(k,k))] \\ &\quad (M+\varepsilon) - \delta u_{12}(k,k)(M^2 - \varepsilon^2) \end{aligned}$$

Given $\varepsilon > 0$, one can find $z^s \to k(z^s \neq k)$ such that

$$[h(k) - h(z^s)]/(k - z^s) \geq (M - \varepsilon) \tag{7.17}$$

One can then find S' and x^s such that $h(x^s) = z^s$ for $s \geq S'$. Then $x^s \to k(x^s \neq k)$ as $s \to \infty$, and one can find $S \geq S'$, such that for $s \geq S$,

$$[h(x^s) - h(k)]/(x^s - k) \leq (M + \varepsilon) \tag{7.18}$$

Using (7.15) with $x = x^s$, (7.17) and (7.18), we have

$$\begin{aligned} \widetilde{u}_{21} &= \left[(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11}) - \delta\overline{u}_{12}\left(\frac{h(z^s)-h(k)}{z^s-k}\right)\right] \\ &\quad \left[\frac{h(x^s)-h(k)}{x^s-x}\right] \\ &\leq [(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11}) - \delta\overline{u}_{12}(M-\varepsilon)]\left[\frac{h(x^s)-h(k)}{x^s-k}\right] \\ &\leq [(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11}) - \delta\overline{u}_{12}(M-\varepsilon)](M+\varepsilon) \end{aligned}$$

Letting $x^s \to k$, we get

$$\begin{aligned} u_{21}(k,k) &\leq [(-u_{22}(k,k)) + \delta(-u_{11}(k,k)) - \delta u_{12}(k,k)(M-\varepsilon)] \\ (M+\varepsilon) &< u_{21}(k,k) \end{aligned}$$

a contradiction, which establishes the claim.

Claim 2: $g(M) < 0$ is not possible. Otherwise,

$$u_{21}(k,k) + [u_{22}(k,k) + \delta u_{11}(k,k)]M + \delta u_{12}(k,k)M^2 < 0$$

We can choose $\varepsilon > 0$ such that

$$\begin{aligned} u_{21}(k,k) &< [(-u_{22}(k,k)) + \delta(-u_{11}(k,k))](M-\varepsilon) - \delta u_{12}(k,k) \\ &\quad (M^2 - \varepsilon^2) \end{aligned}$$

One can find $x^s \to k(x^s \neq k)$ such that

$$[h(k) - h(x^s)]/(k - x^s) \geq (M - \varepsilon) \tag{7.19}$$

Define $z^s = h(x^s)$. Then $z^s \to k(z^s \neq k)$ and we can find S, such that for $s \geq S$,

$$[h(k) - h(z^s)]/(k - z^s) \leq (M + \varepsilon) \tag{7.20}$$

Using (7.15) with $x = x^s$, (7.19) and (7.20), we have

$$\begin{aligned} \widetilde{u}_{21} &= \left[(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11}) - \delta\overline{u}_{12}\left[\frac{h(k)-h(z^s)}{k-z^s}\right]\right] \\ &\quad \left[\frac{h(x^s)-h(k)}{x^s-k}\right] \\ &\geq [(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11}) - \delta\overline{u}_{12}(M+\varepsilon)](M-\varepsilon) \end{aligned}$$

Letting $x^s \to k$, we get

$$\begin{aligned} u_{21}(k,k) &\geq [(-u_{22}(k,k) + \delta(-u_{11}(k,k))](M-\varepsilon) \\ &\quad -\delta u_{12}(k,k)(M^2 - \varepsilon^2) \\ &> u_{21}(k,k), \end{aligned}$$

a contradiction. From Claims 1 and 2, we get $g(M) = 0$, so $M = \lambda_1$ or λ_2.

Step 4: Let $x \in N(k)$, with $x \neq k$. To be precise, let $x < k$, so that $h(x) < k$ and $h^2(x) < k$. [The case $x > k$ can be handled similarly.] Then $V'(x) = u_1(x, h(x))$ and $\delta V'(h(x)) = \delta u_1(h(x), h^2(x)) = -u_2(x, h(x))$. Similarly $V'(k) = u_1(k, k)$ and $\delta V'(k) = \delta u_1(k, k) = -u_2(k, k)$. Thus, using concavity of u, we obtain the following two inequalities:

$$\begin{aligned} u(x, h(x)) + \delta V'(h(x))h(x) - V'(x)x &\geq u(k, k) + \delta V'(h(x))k \\ &\quad -V'(x)k \\ u(k, k) + \delta V'(k)k - V'(k)k &\geq u(x, h(x)) + \delta V'(k)h(x) \\ &\quad -V'(k)x \end{aligned}$$

Adding the inequalities and transposing terms

$$\delta[V'(h(x)) - V'(k)][k - h(x)] \leq [V'(x) - V'(k)][k - x]$$

Iterating on this relationship, we get

$$\delta[V'(h(x)) - V'(k)][k - h^2(x)] \leq [V'(x) - V'(k)][k - x]$$

This yields the inequality

$$\delta^2[k - h^2(x)]/[k - x] \leq [V'(x) - V'(k)]/[V'(h^2(x) - V'(k)] \tag{7.21}$$

We claim now that

$$\delta^2[k - h^2(x)]/[k - x] \leq 1 \tag{7.22}$$

For, if (7.22) were violated, we would have $[k - h^2(x)] > [k - x]/\delta^2 > [k - x]$. Thus, we must have $h^2(x) < x < k$, and $V'(h^2(x)) \geq V'(x) \geq V'(k)$, so that $[V'(x) - V'(k)]/[V'(h^2(x)) - V'(k)] \leq 1$. But then by using (7.21), (7.22) must hold, a contradiction. This establishes (7.22).

Step 5: By the analysis in Step 3, we know that M is either λ_1 or λ_2 and m is either λ_1 or λ_2. We now show that $M = m = \lambda_1$.

By (7.15), we have

$$\widetilde{u}_{21} + [(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11})]\{[h(x) - h(k)]/(k - x)\} +$$
$$\delta\overline{u}_{12}\{[h^2(k) - h^2(x)]/(k - x)\} = 0$$

Using (7.22) we get

$$[(-\widetilde{u}_{22}) + \delta(-\overline{u}_{11})]\{[h(x) - h(k)]/(x - k)\} \leq \widetilde{u}_{21} + [\overline{u}_{12}/\delta]$$

Thus, we obtain (by letting $x \to k$),

$$|Dh(k)| \leq (1 + \delta)[+u_{12}(k,k)]/\delta[(-u_{22}(k,k)) + \delta(-u_{11}(k,k))]$$

for every Dini-derivative at k. Using C1 we get $|Dh(k)| < (1/\delta) < \lambda_2$. Thus, $M < \lambda_2$ and consequently $M = \lambda_1 < 1$. Since m is either λ_1 or λ_2, and $m \leq M$, $m = \lambda_1$, also.

The above analysis implies that h is differentiable at k, and $h'(k) = \lambda_1$.

We summarize our finding in the following result.

Proposition 7.3: *Under C1, the policy function h is differentiable at k, and*

$$|h'(k)| = |\lambda_1|$$

where λ_1 is the characteristic root associated with the Ramsey-Euler equation at (k,k), with the smaller absolute value.

8. Uniqueness of the Stationary Optimal Stock

In general, a non-trivial stationary optimal stock need not be unique. However, one can provide a useful sufficient condition on the reduced-form utility function under which a uniqueness result can be obtained, independent of the discount factor. These two topics are discussed in Sections 8.1 and 8.2.

An especially useful theory of uniqueness can be developed when (Ω, u) are given, and the discount factor is restricted to take values close to unity and we examine below the elements of this theory in some detail in Section 8.3.

8.1 An Example of Non-Uniqueness

We examine a situation in the context of Example 3 of Section 2 by specifying (w, f, δ), for which we can verify the existence of two non-trivial stationary optimal stocks.

Let $X = [0, 4]$, and the function $f : X \to X$ be specified by

$$f(x) = 2x^{1/2} \text{ for all } x \in X$$

We show now how two stocks can be shown to be stationary optimal stocks by appropriate choice of (w, δ).

Let us pick $x' = (1/4)$ and $x'' = (1/9)$. Then $f(x') = 2(x')^{1/2} = 1$ and $f(x'') = 2(x'')^{1/2} = 2/3$. Thus, consumption levels associated with stationary programs $(x')_0^\infty$ and $(x'')_0^\infty$ are $c' = (3/4)$ and $c'' = (5/9)$.

In order that these stocks be stationary optimal stocks, it is enough to check that the Ramsey-Euler equations hold for our choice of (w, δ). To make computations easier, we take w to be of the additive separable form:

$$w(c, x) = g(c) + h(x)$$

Then the Ramsey-Euler equations would hold at (x', c') and (x'', c'') if

$$\begin{aligned} g'(c') &= \delta[g'(c')f'(x') + h'(x')] \\ g'(c'') &= \delta[g'(c'')f'(x'') + h'(x'')] \end{aligned}$$

These can be rewritten as:

$$\begin{aligned} g'(c')[1 - \delta f'(x')] &= \delta h'(x') \\ g'(c'')[1 - \delta f'(x'')] &= \delta h'(x'') \end{aligned}$$

Since $f'(x) = [1/x^{1/2}]$, we have $f'(x') = 2$ and $f'(x'') = 3$. Let us choose $\delta = (1/6)$. Then $\delta f'(x') = 1/3$ and $\delta f'(x'') = 1/2$, so that $[1-\delta f'(x')] = (2/3)$ and $[1-\delta f'(x'')] = (1/2)$. Thus, in order for the Ramsey-Euler equations to hold, all we need to ensure is that

$$\begin{aligned} g'(c') &= (3/2)(1/6)h'(x') \\ g'(c'') &= 2(1/6)h'(x'') \end{aligned}$$

We can clearly simplify our calculations by taking h to be linear; so, choose $h(x) = 12x$ for $x \in X$. Then, our restrictions on g are given by

$$\begin{aligned} g'(c') &= 3 \\ g'(c'') &= 4 \end{aligned}$$

So, all one needs is to provide an increasing, concave function g on X with these two-point restrictions. Clearly, a two-parameter function will suffice for our purpose, and we specify g to be the quadratic map:

$$g(c) = ac - bc^2 \text{ for } c \in X$$

Then we need (a, b) to satisfy

$$\begin{aligned} a - 2bc' &= a - 2b(3/4) = 3 \\ a - 2bc'' &= a - 2b(5/9) = 4 \end{aligned}$$

This yields $a = (453/144)$ and $b = (7/72)$. It can be checked that g is strictly increasing and strictly concave on X.

To summarize, with $f(x) = 2x^{1/2}$ for $x \in X$, $w(c, z) = (ac - bc^2) + 12x$ for $(c, x) \in X^2$ (where $a = (453/144)$ and $b = (7/72)$), and $\delta = (1/6)$, both $x' = (1/4)$ and $x'' = (1/9)$ are non-trivial stationary optimal stocks.

8.2 A Sufficient Condition for Uniqueness

Let us assume that we are given a dynamic optimization model (Ω, u, δ) for which u is twice contnuously differentiable on int Ω, with $u_{11}(x, x) < 0$ and $u_{22}(x, x) \leq 0$ for all $(x, x) \in$ int Ω.

If x' and x'' are two distinct stationary optimal stocks, with $x'' > x'$, such that $(x', x') \in \text{int}\ \Omega$ and $(x'', x'') \in \text{int}\ \Omega$, then the Ramsey-Euler equations yield:

$$u_2(x', x') + \delta u_1(x', x') = 0 \tag{8.1}$$

$$u_2(x'', x'') + \delta u_1(x'', x'') = 0 \tag{8.2}$$

Using the Mean-Value theorem, we can find z and y in (x', x''), such that

$$u_2(x', x') - u_2(x'', x'') = [u_{21}(z, z) + u_{22}(z, z)](x' - x'') \tag{8.3}$$

$$\delta u_1(x', x') - \delta u_1(x'', x'') = [\delta u_{11}(y, y) + \delta u_{12}(y, y)](x' - x'') \tag{8.4}$$

Then $(y, y) \in \text{int}\ \Omega$ and $(z, z) \in \text{int}\ \Omega$, so $u_{11}(y, y) < 0$ and $u_{22}(z, z) < 0$.

Let us assume that the following condition holds:

$$u_{12}(x, x) \leq 0 \text{ for all } (x, x) \in \text{ int } \Omega \tag{8.5}$$

Then adding (8.3) and (8.4), we would get

$$\begin{aligned} 0 = (x'' - x')[(-u_{21}(z, z)) + \\ (-u_{22}(z, z)) + \delta(-u_{11}(y, y)) + \\ \delta(-u_{12}(y, y))] > 0 \end{aligned}$$

a contradiction. Thus, (8.5) is sufficient to ensure uniqueness.

Remark 8.1: (i) In Example 4 of Section 2, we have $u_{12}(x, z) < 0$ for all x, z in $(0, 1) \times (0, 1)$. Thus, (8.5) is satisfied, and a stationary optimal stock must be unique.

(ii) If $u_{12}(x, z) < 0$ for (x, z) in int Ω, then we know that the policy function, h, is decreasing in a neighborhood of x^o, if $(x^o, h(x^o)) \in \text{int}\ \Omega$. This rules out multiple stationary optimal stocks, x, with (x, x) in the interior of Ω. However, (8.5) is a weaker condition ensuring uniqueness.

(iii) The uniqueness of a non-trivial stationary optimal stock in the one-sector model of optimal growth (Example 1) can be viewed as a special case of the above theory, even though, defining $u(x,z) = w(f(x)-z)$, we have $u_{12}(x,z) = [-w''(f(x)-z)]f'(x) > 0$.

To see this, note that if there are two non-trivial stationary optimal stocks, x', x'', then x' must solve

$$\left.\begin{array}{c}\text{Max } w(f(x)-z) \\ \text{subject to } \delta f(x) - z \geq \delta f(x') - x' \\ (x,z) \in \Omega\end{array}\right\} (P')$$

and x'' must solve

$$\left.\begin{array}{c}\text{Max } w(f(x)-z) \\ \text{subject to } \delta f(x) - z \geq \delta f(x'') - x'' \\ (x,z) \in \Omega\end{array}\right\} (P'')$$

But with w strictly increasing (P') and (P'') are the same as:

$$\left.\begin{array}{c}\text{Max } f(x) - z \\ \text{subject to } \delta f(x) - z \geq \delta f(x') - x' \\ (x,z) \in \Omega\end{array}\right\} (P_1')$$

and

$$\left.\begin{array}{c}\text{Max } f(x) - z \\ \text{subject to } \delta f(x) - z \geq \delta f(x'') - x'' \\ (x,z) \in \Omega\end{array}\right\} (P_1'')$$

Thus, x' solves (P_1') and x'' solves (P_1''). Then, defining $\mathcal{U}(x,z) = f(x) - z$, x' and x'' are non-trivial stationary optimal stocks for the model $(\Omega, \mathcal{U}, \delta)$. But since $\mathcal{U}_{12} = 0$, $\mathcal{U}$ satisfies (8.5) and with $f'' > 0$ on *int* $X, \mathcal{U}_{11} < 0$ and $\mathcal{U}_{22} = 0$, so we must have uniqueness of the stationary optimal stock; that is, x' must be equal to x''.

8.3 Uniqueness of the Stationary Optimal Stock for High Discount Factors

Let (Ω, u) be given satisfying A1, A2, and A5. Consider the maximization problem:

$$\left.\begin{array}{c} \text{Max } u(x,z) \\ \text{subject to } z - x \geq 0 \\ (x,z) \in \Omega \end{array}\right\} (P)$$

Clearly, there is a solution $(\overline{x}, \overline{z})$ to (P), and so $(\overline{x}, \overline{x})$ solves (P) as well. We refer to $\overline{x}$ as a *golden-rule stock.*

Assume, now, that the following condition holds:

C2: (i) $(\overline{x}, \overline{x}) \in \text{int } \Omega$

(ii) u is twice continuously differentiable on int Ω, with the Hessian of u negative definite at $(\overline{x}, \overline{x})$.

Note that there is a neighborhood N (in *int* Ω) containing $(\overline{x}, \overline{x})$, such that the Hessian of u is negative definite on N.

It follows that $(\overline{x}, \overline{x})$ is the only solution to (P). For if (x, z) were any solution to (P), then so is (x, x), and by convexity of Ω and concavity of u, so is $(\lambda x + (1 - \lambda)\overline{x},\ \lambda x + (1 - \lambda)\overline{x})$ for all $0 \leq \lambda \leq 1$. For $0 < \lambda < 1$ and sufficiently close to zero, this is in N, but since u is strictly concave on N, we must have $x = \overline{x}$. Similarly, using the fact that (z, z) solves (P), we can infer that $z = \overline{x}$. Thus $(x, z) = (\overline{x}, \overline{x})$.

We can now find $\varepsilon_1 > 0$ such that $(x, x) \in \text{int } \Omega$ for all $x \in [\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1] \equiv I_1 \subset N$.

Since the Hessian of u is negative definite on N, we have

$$\alpha = -[u_{21}(\overline{x}, \overline{x}) + u_{22}(\overline{x}, \overline{x}) + u_{11}(\overline{x}, \overline{x}) + u_{12}(\overline{x}, \overline{x})] > 0$$

Thus, there is $0 < \varepsilon_2 < \varepsilon_1$ such that for all x in $I_2 \equiv [\overline{x} - \varepsilon_2, \overline{x} + \varepsilon_2]$, we have

$$-[u_{21}(x, x) + u_{22}(x, x) + u_{11}(x, x) + u_{12}(x, x)] > (\alpha/2)$$

We can now find $0 < \rho_1 < 1$ such that for all $\rho_1 < \rho < 1$ and all $x \in I_2$,

$$-[u_{21}(x, x) + u_{22}(x, x) + \rho u_{11}(x, x) + \rho u_{12}(x, x)] > (\alpha/4)$$

We can find $0 < \beta < [u(\overline{x},\overline{x}) - u(0,0)]$ such that whenever $(x,x) \in \Omega$ and $|x - \overline{x}| \geq \varepsilon_2$, we have $u(x,x) \leq u(\overline{x},\overline{x}) - \beta$. Now, pick $0 < \varepsilon_3 < \varepsilon_2$ such that $(\overline{x}, \overline{x} + \varepsilon_3) \in \text{int } \Omega$ and $u(\overline{x}, \overline{x} + \varepsilon_3) > u(\overline{x},\overline{x}) - \beta$. Next, find $\rho_1 < \rho_2 < 1$ such that for $\rho_2 < \rho < 1$, we have $\rho(\overline{x} + \varepsilon_3) - \overline{x} > 0$.

Now, choose the discount factor, δ, in $(\rho_2, 1)$. Since $(\overline{x}, \overline{x} + \varepsilon_3) \in \Omega$, with $\delta(\overline{x} + \varepsilon_3) > \overline{x}$, we have $(\overline{x} + \varepsilon_3) > \overline{x}/\delta$ and so $(\overline{x}, \overline{x}/\delta) \in \Omega$, and since $u(\overline{x}, \overline{x} + \varepsilon_3) > u(\overline{x},\overline{x}) - \beta > u(0,0)$, we have $u(\overline{x}, \overline{x}/\delta) > u(0,0)$ and so (Ω, u, δ) is δ-normal. Thus, there exists a discounted golden-rule stock. This is also a non-trivial stationary optimal stock.

We show now that there can be at most one non-trivial stationary stock (SOS), k. We note that if k is a non-trivial SOS, then $(k,k) \in \text{int } \Omega$. To see this note that since $(\overline{x},\overline{x}) \in \text{int } \Omega$, all $0 < x \leq \overline{x}$ must satisfy $(x,x) \in \text{int } \Omega$ as well. Thus, if k' is a non-trivial SOS such that (k',k') is not in the interior of Ω, then $k' > \overline{x}$. But the program $(k', \overline{x}, \overline{x}, ...)$ from k clearly has a higher discounted sum of utilities than $(k', k', k', ...)$, so k' cannot be an SOS.

We know that any SOS, k, with $(k,k) \in \text{int } \Omega$ is a discounted golden-rule stock, and so (k,k) must solve the problem:

$$\begin{aligned} &\text{Max } u(x,z) \\ &\text{subject to } \delta z - x \geq \delta k - k \\ &(x,z) \in \Omega \end{aligned}$$

Since $\delta(\overline{x} + \varepsilon_3) - \overline{x} > 0 \geq \delta k - k$, we must have $u(k,k) \geq u(\overline{x}, \overline{x} + \varepsilon_3) > u(\overline{x},\overline{x}) - \beta$. It follows then that $|k - \overline{x}| < \varepsilon_3$. Since $0 < \varepsilon_3 < \varepsilon_2$, $k \in I_2$, and since $1 > \delta > \rho_2$, we have $1 > \delta > \rho_1$, and so

$$-[u_{21}(k,k) + u_{22}(k,k) + \delta u_{11}(k,k) + \delta u_{12}(k,k)] > (\alpha/4)$$

Define $I_3 = [\overline{x} - \varepsilon_3, \overline{x} + \varepsilon_3]$ and $G : I_2 \rightarrow \Re$ by

$$G(x) = u_2(x,x) + \delta u_1(x,x)$$

Then $G(k) = 0$ and $G'(k) < 0$.

Let z be any point in I_2 such that $G(z) = 0$. Then $(z, z) \in \text{int}\,\Omega$ and (z, z) satisfies the Ramsey-Euler equation and is a stationary optimal stock. Thus $z \in I_3$ and $G'(z) < 0$ by the above argument.

It suffices then to show that there can be at most one zero of G in I_2. To this end, define $F : I_2 \to \Re$ by:

$$F(z) = \int_{\overline{x}-\varepsilon_2}^{z} G(x)dx$$

Then F is continuous on I_2 and twice continuously differentiable on int I_2 with $F'(z) = G(z)$ and $F''(z) = G'(z)$ for all $z \in$ int I_2. Further F is quasi-concave on I_2. Otherwise there would exist x, x' in I_2 with $x < x'$, and $0 < \lambda < 1$, such that $F[\lambda x + (1 - \lambda)x'] < \min[F(x), F(x')]$. Then F attains an interior minimum at some z in $[x, x']$, so that $F'(z) = 0$ and $F''(z) \geq 0$. But then $z \in$ int I_2 and $G(z) = 0$ and $G'(z) \geq 0$, a contradiction.

Suppose a and b are distinct zeroes of G in I_2. Then a, $b \in I_3$ and $G(z) = G(b) = 0$ and $G'(a) < 0$ and $G'(b) < 0$. Thus, $F'(a) = F'(b) = 0$, and $F''(a) < 0$ and $F''(b) < 0$. Since F is C^2 on int I_2, we can find a neighborhood $N(a)$ of a, such that for all $x \in N(a)$, $F''(x) < 0$. Then one can find $0 < \lambda < 1$ such that $[\lambda a + (1-\lambda)b] \in N(a)$, and by the strict concavity of F on $N(a)$, we must have $F[\lambda a + (1 - \lambda)b] - F(a) < F'(a)[\lambda a + (1 - \lambda)b - a] = 0$, so that $F[\lambda a + (1 - \lambda)b] < F(a)$. By the quasi-concavity of F, we have $F[\lambda a + (1 - \lambda)b] > \min[F(a), F(b)]$. Thus $F(a) > F(b)$. Exchanging the roles of a and b in the above argument, $F(b) > F(a)$, a contradiction. Thus, there is at most one zero of G in I_2.

9. Global Asymptotic Stability of the Stationary Optimal Stock

Even when there is a unique non-trivial stationary optimal stock, it need not be globally asymptotically stable; that is, optimal programs from other inital stocks need not converge to the SOS over time. This topic is discussed in Section 9.1.

Parallel to the theory of uniqueness developed in Section 8.3, when (Ω, u) are given and the discount factor, δ, is restricted to take values close to unity, then a theory of global asymptotic stability of the (unique) non-trivial stationary optimal stock can be obtained. This topic (developed in Section 9.2), which falls under the general subject matter of "turnpike theory", is one of the major developments in dynamic optimization theory. It relies on practically all the material developed in the previous sections, and the argument is subtle. In particular, it is extremely important to obtain the relevant bound on the discount factor in terms of conditions (which can be checked easily) placed solely on (Ω, u).

Our approach to this topic is deliberately one-dimensional, although it uses many ideas which are used in the literature on turnpike theory in the multidimensional setting. We show that for discount factors close to unity, there are no period-two optimal programs, and global asymptotic stability of the non-trivial SOS is a consequence of this fact (in our one-dimensional state variable setting).

This has two advantages. First, it clarifies the relation between turnpike theory and the presence/absence of period-two optimal cycles, an idea which is especially useful as a backdrop in the study of periodic and chaotic optimal programs. Second, the problem of ruling out period-two optimal cycles (when the discount factor is close to unity) is simply an extension of the problem of ruling out multiple stationary optimal stocks. Thus, the idea that is exploited is that a stronger bound on the discount factor than was used to generate uniqueness of the non-trivial stationary optimal stock also ensures global asymptotic stability of this SOS.

9.1 A Counterexample to the Turnpike Property

We examine a situation in the context of Example 2.4, with $\alpha = \beta = 0.5$, the case examined by Weitzman. Thus $X = [0, 1]$, $\Omega = X^2$ and $u : \Omega \to \Re$ is given by

$$u(x, z) = x^{1/2}(1 - z)^{1/2} \tag{9.1}$$

Given any $x \in (0, 1)$, we can show that there is $z \in (0, 1)$ such that $(x, z, x, z, ...)$ is an optimal program from x. To see this simply consider the Ramsey-Euler equation

$$u_2(x,z) + \delta u_1(z,x) = 0 \tag{9.2}$$

and note that (9.1) yields

$$\frac{z}{(1-z)} = \delta^2 \frac{(1-x)}{x} \tag{9.3}$$

Denoting by y the right hand side of (9.3), we note that $y > 0$ and

$$z = y/(1+y) \tag{9.4}$$

leading to a unique choice of z in $(0,1)$, given x in $(0,1)$, to (9.2). Further, with $x \in (0,1)$ and with z defined by (9.4), that is,

$$z = \delta^2(1-x)/[x + \delta^2(1-x)] \tag{9.5}$$

we can also check that

$$u_2(z,x) + \delta u_1(x,z) = 0 \tag{9.6}$$

To see this, use (9.1) in (9.6) to get

$$\frac{z}{(1-z)} = \delta^2 \frac{(1-x)}{x}$$

which is exactly the same equation as (9.3).

Thus, for $x \in (0,1)$, and z defined by (9.5), we have (9.2) and (9.6) holding. Thus $(x, z, x, z, ...)$ is the optimal program from x.

Further, $k = \delta/(1+\delta)$ is the unique stationary optimal stock. And, for $x \in (0,k)$, $z = \delta^2(1-x)/[x+\delta^2(1-x)] > k$, while for $x \in (k,1)$, $z < k$. Thus, for every $x \in (0,1)$, $x \neq k$, there is a period two optimal program. There is, thus, no tendency for optimal programs from $x \neq k$ to converge to the stationary optimal stock, k.

Remark 9.1: The above analysis also demonstrates that the (optimal) policy function, h, for the dynamic optimization model (Ω, u, δ) can be solved explicitly. In fact it is given by (9.5).

9.2 A Turnpike Theorem for High Discount Factors

The example (due to Weitzman and discussed in Samuelson (1973)) presented in the previous subsection has the feature that there are period-two optimal cycles from every initial state (in the interior of the state space) other than the (unique) stationary optimal state, *independent of how high the discount factor is.*

While this destroys the hope of proving a general turnpike theorem for high discount factors, it was observed by Samuelson (1973) that if the utility function has a negative definite Hessian at the golden-rule, then for high discount factors, optimal cycles would be ruled out, and stability (at least of the local kind) of the stationary optimal stock would be ensured.

This observation led to a large literature dealing with the turnpike theorem for high discount factors in a variety of settings. Our objective is to carry through a complete analysis of Samuelson's original idea, to establish an appropriate turnpike theorem. That is, a basic ingredient of our approach is that in the situation described by Samuelson, optimal cycles can be ruled out, and this, in turn, ensures global asymptotic stability of the stationary optimal stock.

We briefly describe the basic steps of our analysis before presenting the material formally. Following Samuelson, we assume that the utility function has a negative definite Hessian at the golden-rule. That is, we maintain condition C2 (used in Section 8.3 to ensure uniqueness of the SOS at high discount factors). As a first step, we choose an appropriately small neightborhoood (N) of the golden-rule, on which the utility function has a negative definite Hessian. Anticipating the proof a bit, this neighborhood is so chosen that when the discount factor (to be chosen later) is high, there are no period-two optimal cycles confined to this neighborhood. The idea here basically mimics the analysis of uniqueness of the SOS for high discount factors, with the single Ramsey-Euler equation (characterizing an interior SOS) replaced by a pair of Ramsey-Euler equations (characterizing an interior period-two cycle).

As a second step, we put bounds on the discount factor, so that the following hold:

(i) there is a unique non-zero SOS, k_δ, in the state space X (which can be ensured, by following the analysis in Section 8.3, to be in the neighborhood N).

(ii) if there is a period-two optimal cycle, it must belong to the neighborhood, N.

(iii) condition C1 of Section 7.2 holds ensuring that the policy function, h_δ, satisfies $|h'_\delta(k_\delta)| < 1$.

(iv) if the initial stock $x \in (0, b]$, $(x, 0, 0, ...)$ is not optimal starting from x.

Findings (i) and (ii) of the second step can be used with the first step to ensure there are no period-two optimal cycles when the discount factor is high. Then, a well-known result from one-dimensional dynamics ensures us that if (x_t) is the optimal program from $x \in X$, then x_t converges to one of the fixed points of h_δ.

Now (iii) and (iv) can be used to ensure that x_t cannot converge to zero. Thus, x_t must converge to the unique non-zero SOS, k_δ, and the turnpike theorem is established.

We now turn to the formal details of turnpike theory. As mentioned above, we maintain condition C2, so that $(\overline{x}, \overline{x})$ is the unique golden-rule (solution to problem (P) in Section 8.3).

Step 1: Using the fact that the Hessian of u is negative definite at $(\overline{x}, \overline{x})$, we can find $\alpha > 0$ such that

$$(-u_{11}(\overline{x}, \overline{x})) + (-u_{22}(\overline{x}, \overline{x})) > \alpha + 2|u_{12}(\overline{x}, \overline{x})| \tag{9.7}$$

Given that $(\overline{x}, \overline{x}) \in \text{int } \Omega$, we can find $0 < \varepsilon_1 < \overline{x}$, such that $N_1(\overline{x}) \equiv [\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1]$ is in the interior of X, and if $(x, z) \in N_1(\overline{x})^2$, then $(x, z) \in \text{int } \Omega$, and the Hessian of u is negative definite on $N_1(\overline{x})^2$.

Next, given (9.7), we can find $0 < \varepsilon_2 < \varepsilon_1$, such that $N_2(\overline{x}) \equiv (\overline{x} - \varepsilon_2, \overline{x} + \varepsilon_2) \subset N_1(\overline{x})$, and for all $(x, z) \in N_2(\overline{x})^2$,

$$\det \begin{bmatrix} u_{21}(x, z) + u_{12}(z, x) & u_{22}(x, z) + u_{11}(z, x) \\ u_{22}(x, z) + u_{11}(x, z) & u_{21}(z, x) + u_{12}(x, z) \end{bmatrix} < -(\alpha/2) \tag{9.8}$$

We claim now that we can find $\widehat{\beta} > 0$, such that if $((x, z), (z, x)) \in \Omega^2$ and (x, z) is not in $N_2(\overline{x})^2$, then

$$u(x, z) + u(z, x) < 2u(\overline{x}, \overline{x}) - \widehat{\beta} \tag{9.9}$$

Otherwise, we can find a convergent sequence $(x^s, z^s)_{s=1}^{\infty}$ converging to (x^0, z^0), with $((x^s, z^s), (z^s, x^s))$ in Ω^2 for $s = 1, 2, 3, ...$, and

$$u(x^s, z^s) + u(z^s, x^s) \geq 2u(\overline{x}, \overline{x}) - (1/s) \tag{9.10}$$

Since $(\overline{x}, \overline{x})$ belongs to the interior of $N_2(\overline{x})^2$, we can find $0 < \lambda < 1$ such that $(\widetilde{x}^s, \widetilde{z}^s)$ and $(\widetilde{z}^s, \widetilde{x}^s)$ belongs to $N_3(\overline{x})^2$ for $s = 1, 2, ...,$ where $N_3(\overline{x}) = [\overline{x} - 0.5\varepsilon_2, \overline{x} + 0.5\varepsilon_2]$, and

$$(\widetilde{x}^s, \widetilde{z}^s) = \lambda(x^s, z^s) + (1 - \lambda)(\overline{x}, \overline{x}) \tag{9.11}$$

Clearly $(\widetilde{x}^s, \widetilde{z}^s)$ converges to $(\widetilde{x}^0, \widetilde{z}^0) \equiv [\lambda(x^0, z^0) + (1 - \lambda)(\overline{x}, \overline{x})]$, and $(\widetilde{x}^0, \widetilde{z}^0) \in N_3(\overline{x})^2$.

Using (9.10) and (9.11), we have for $s = 1, 2, 3, ...$

$$u(\widetilde{x}^s, \widetilde{z}^s) + u(\widetilde{z}^s, \widetilde{x}^s) \geq 2u(\overline{x}, \overline{x}) - (\lambda/s) \tag{9.12}$$

Letting $s \longrightarrow \infty$, we get

$$u(\widetilde{x}^0, \widetilde{z}^0) + u(\widetilde{z}^0, \widetilde{x}^0) \geq 2(\overline{x}, \overline{x}) \tag{9.13}$$

Note that $(\widetilde{x}^0, \widetilde{z}^0) \neq (\overline{x}, \overline{x})$. Thus, if $\widetilde{x}^0 = \widetilde{z}^0$, (9.13) contradicts the uniqueness of the golden-rule. And if $\widetilde{x}^0 \neq \widetilde{z}^0$, then using the fact that $(0.5\widetilde{x}^0 + 0.5\widetilde{z}^0, 0.5\widetilde{z}^0 + 0.5\widetilde{x}^0)$ is in $N_2(\overline{x})^2$ and u is strictly concave on $N_2(\overline{x})^2$, we get from (9.13)

$$\begin{aligned} u(0.5\widetilde{x}^0 + 0.5\widetilde{z}^0, 0.5\widetilde{z}^0 + 0.5\widetilde{x}^0) &> \\ 0.5u(\widetilde{x}^0, \widetilde{z}^0) + 0.5u(\widetilde{z}^0, \widetilde{x}^0) &\geq u(\overline{x}, \overline{x}) \end{aligned}$$

which clearly contradicts the definition of the golden-rule. This establishes (9.9).

Define $\beta = \min[\widehat{\beta}, u(\overline{x}, \overline{x}) - u(0, 0)]$, and note that $\beta > 0$. Pick $0 < \varepsilon_3 < \varepsilon_2$ such that

$$u(\overline{x}, \overline{x} + \varepsilon_3) > u(\overline{x}, \overline{x}) - (\beta/8). \tag{9.14}$$

Step 2: We now put suitable bounds on the range of the discount factor.

First, using (9.8), we can find $0 < \rho_1 < 1$ such that for all $\rho_1 < \rho < 1$ and all $(x,z) \in N_2(\overline{x})^2$,

$$\det \begin{bmatrix} u_{21}(x,z) + \rho u_{12}(z,x) & u_{22}(x,z) + \rho u_{11}(z,x) \\ u_{22}(z,x) + \rho u_{11}(x,z) & u_{21}(z,x) + \rho u_{12}(x,z) \end{bmatrix} < -(\alpha/4) \tag{9.15}$$

Second, given (9.9) and the definition of β, we can find $\rho_1 < \rho_2 < 1$ such that for all $((x,z),(z,x)) \in \Omega^2$ with (x,z) not in $N_2(\overline{x})^2$, and all $\rho_2 < \rho < 1$, we have

$$u(x,z) + \rho u(z,x) < u(\overline{x},\overline{x}) + \rho u(\overline{x},\overline{x}) - (\beta/2). \tag{9.16}$$

Now, given (9.14), we have for $\rho_2 < \rho < 1$, $(\beta/8) + \rho u(\overline{x},\overline{x}+\varepsilon_3) > \rho(\beta/8)+\rho u(\overline{x},\overline{x}+\varepsilon_3) > \rho u(\overline{x},\overline{x})$. Therefore, for all $((x,z),(z,x)) \in \Omega^2$ with (x,z) not in $N_2(\overline{x})^2$, and all $\rho_2 < \rho < 1$,

$$\begin{aligned} u(x,z) + \rho u(z,x) \; &< \; u(\overline{x},\overline{x}+\varepsilon_3) + (\beta/8) \\ &\quad +\rho u(\overline{x},\overline{x}+\varepsilon_3) + (\beta/8) - (\beta/2) \end{aligned}$$

so that

$$u(x,z) + \rho(z,x) < u(\overline{x},\overline{x}+\varepsilon_3) + \rho u(\overline{x},\overline{x}+\varepsilon_3) - (\beta/4) \tag{9.17}$$

Finally, we pick $\rho_2 < \rho_3 < 1$, such that for all $\rho_3 \leq \rho < 1$, we have

$$\rho \frac{[u(\overline{x},\overline{x}) - u(0,0)]}{(1-\rho)} > [u(b,0) - u(\overline{x}-\varepsilon_2,\overline{x})] \tag{9.18}$$

and

$$\rho^2(\overline{x}+\varepsilon_3) > \overline{x} \tag{9.19}$$

Proposition 9.1: *Let (Ω, u) be given satisfying condition C2. If the discount factor δ satisfies $\rho_3 \leq \delta < 1$, an optimal program from $x \in X$ cannot exhibit a period-two cycle.*

Proof: Given (Ω, u), let ρ be a number in $[\rho_3, 1)$. Suppose for (Ω, u, ρ), there is an optimal program from $x' \in X$, such that it exhibits a period-two cycle. Then, there is $z' \in X$, $z' \neq x'$, such that $(x', z', x', z', ...)$ is optimal from x'. Without loss of generality, we can take $x' > z'$. Then $x' > 0$ and $(z', x') \in \Omega$ with $x' > 0$, so we can find $p \in \Re_+$ such that

$$\begin{aligned} & u(x', z') + \rho u(z', x') + p^2 px' - px' \\ \geq \; & u(x'', z'') + \rho u(z'', y'') + \rho^2 py' - px' \end{aligned} \tag{9.20}$$

for all $((x'', z'')(z'', y'')) \in \Omega^2$.

Choosing $(x'', z'') = (\overline{x}, \overline{x})$ and $(z'', y'') = (\overline{x}, \overline{x} + \varepsilon_3)$ in (9.20), we get

$$u(x', z') + \rho u(z', x') \geq u(\overline{x}, \overline{x}) + \rho u(\overline{x}, \overline{x} + \varepsilon_3) \tag{9.21}$$

since $[\rho^2 px' - px'] \leq 0$, and $\rho^2 p((\overline{x} + \varepsilon_3) - p\overline{x} \geq 0$ by (9.19) and the fact that $\rho_3 \leq \rho < 1$. Using (9.21) and (9.17) and $u(\overline{x}, \overline{x}) \geq u(\overline{x}, \overline{x} + \varepsilon_3)$, we see that $(x', z') \in N_2(\overline{x})^2$.

Using (9.14), (9.19) and the definition of β, (Ω, u, ρ) is ρ-normal for $\rho_3 \leq \rho < 1$, so there exists a non-trivial SOS, k_ρ. By the analysis of Section 8.3, k_ρ is the unique non-trivial SOS, and indeed k_ρ is the only non-zero SOS. Further, $k_\rho \in N_2(\overline{x})$.

Let $N(\overline{x}) \equiv [\overline{x} - \varepsilon_2, \overline{x} + \varepsilon_2]$ and define the function, $H : N(\overline{x})^2 \times [0, 1] \to \Re$ by

$$H(x, z; \lambda) = \begin{bmatrix} u_2(x, z) + (\lambda\rho_3 + (1 - \lambda))u_1(z, x) \\ u_2(z, x) + (\lambda\rho_3 + (1 - \lambda))u_1(x, z) \end{bmatrix} \tag{9.22}$$

We will show that (i) degree $H(N_2(\overline{x})^2; 0) = -1$, and (ii) degree $H(N_2(\overline{x})^2; \lambda) =$ degree $H(N_2(\overline{x})^2; 0)$ for all λ in $(0, 1]$.

Notice that (9.22) yields:

$$H(x, z; 0) = \begin{bmatrix} u_2(x, z) + u_1(z, x) \\ u_2(z, x) + u_1(x, z) \end{bmatrix} \tag{9.23}$$

Clearly since the golden-rule $(\overline{x}, \overline{x})$ solves problem (P) of Section 8.3, and $(\overline{x}, \overline{x}) \in \text{int } N(\overline{x})^2$, we have $H(\overline{x}, \overline{x}; 0) = 0$. Also, if $(x, z) \in$

$N(\overline{x})^2$ with $(x, z) \neq (\overline{x}, \overline{x})$ satisfies $H(x, z; 0) = 0$, then

$$\begin{aligned} & u(\overline{x}, \overline{x}) + u(\overline{x}, \overline{x}) - u(x, z) - u(z, x) \\ \leq \;& u_1(x, z)(\overline{x} - x) + u_2(x, z)(\overline{x} - z) + u_1(z, x)(\overline{x} - z) \\ & + u_2(z, x)(\overline{x} - x) \\ = \;& 0 \end{aligned}$$

so that $[0.5u(x, z) + 0.5u(z, x)] \geq u(\overline{x}, \overline{x})$. If $x = z$ this contradicts the uniqueness of the golden-rule. If $x \neq z$, then since $[0.5x + 0.5z, 0.5z + 0.5x] \in N(\overline{x})^2$, and u is strictly concave on $N(\overline{x})^2$,

$$\begin{aligned} & u(0.5x + 0.5z, 0.5z + 0.5x) \\ > \;& 0.5u(x, z) + 0.5u(z, x) \geq u(\overline{x}, \overline{x}) \end{aligned}$$

which contradicts the definition of the golden-rule. Thus, $(x, z) = (\overline{x}, \overline{x})$ is the only solution to $H(x, z; 0) = 0$ in $N(\overline{x})^2$. Using (9.8), we have degree of $H(N_2(\overline{x})^2; 0) = -1$.

Now for any λ in $(0, 1]$, defining $\rho = \lambda \rho_3 + (1 - \lambda)$, we have $\rho_2 < \rho_3 \leq \rho < 1$. If (x, z) in $N(\overline{x})^2$ is a solution to $H(x, z; \lambda) = 0$ which is on the boundary of $N(\overline{x})^2$, then

$$\begin{aligned} u_2(x, z) + \rho u_1(z, x) &= 0 \\ u_2(z, x) + \rho u_1(x, z) &= 0 \end{aligned}$$

and so either (i) $x = z$, and x is a stationary optimal stock for (Ω, u, ρ); or (ii) $x \neq z$, and (x, z) is a period-two cycle for (Ω, u, ρ). However, as verified above, in case (i), $x = k_\rho \in N_2(\overline{x})$ and in case (ii), $(x, z) \in N_2(x)^2$. Thus, there is no solution (x, z) to $H(x, z; \lambda) = 0$ on the boundary of $N(\overline{x})^2$. It follows that the degree of $H(N_2(\overline{x})^2; \lambda) = H(N_2(\overline{x})^2; 0) = (-1)$ for all $\lambda \in (0, 1]$.

For any solution (x, z) in $N_2(\overline{x})^2$ to the equation $H((x, z); \lambda) = 0$ we have an index equal to (-1) by (9.15). Thus, there is exactly one solution to the equation $H((x, z); \lambda) = 0$ in $N_2(\overline{x})^2$ for each λ in $[0, 1]$.

Now, let (Ω, u, δ) be a dynamic optimization model with $\rho_3 \leq \delta < 1$. Suppose that (Ω, u, δ) has an optimal program exhibiting a period two-cycle, then there is (x', z') with $x' \neq z'$, such that $(x', z', x', z', ...)$ is optimal from x'. Then, as verified above, $(x', z') \in N_2(\overline{x})^2$. But, there is a stationary optimal stock, k_δ, with $(k_\delta, k_\delta) \in N_2(\overline{x})^2$. Defining $\lambda = (1-\delta)/(1-\rho_3)$, we have λ in $(0, 1]$, and multiple solutions to the equation $H((x, z); \lambda) = 0$ in $N_2(\overline{x})^2$ a contradiction, which establishes the result. Q.E.D.

Proposition 9.2: *Let (Ω, u) be given satisfying condition C2. For every discount factor δ satisfying $\rho_3 \leq \delta < 1$, the optimal policy fuction, h_δ, and non-zero SOS, k_δ, satisfy (i) $h_\delta(x) > x$ for $x \in (0, k_\delta)$; (ii) $h_\delta(x) < x$ for $x \in (k_\delta, b]$.*

Proof: We will write h instead of h_δ to simplify notation. We know that there is a unique non-zero SOS, k_δ, which we denote by k, and $k \in N_2(\overline{x})$.

We claim that for $x \in (0, k)$, $h(x) > x$. Otherwise by uniqueness of non-zero SOS, $h(x) < x$ for all $x \in (0, k)$.

Pick $x \in N_2(\overline{x})$, with $x \in (0, k)$ so that $h(x) \in N_2(\overline{x})$ as well. Then, we have

$$\begin{aligned} -u_2(x, h(x)) &= V'(h(x)) \\ -u_2(k, h(k)) &= V'(h(k)) \end{aligned}$$

Thus, using the Mean-Value theorem,

$$(-\widetilde{u}_{21})(k - x) + (-\widetilde{u}_{22})(k - h(x)) = V'(k) - V'(h(x))$$

where $\widetilde{u}_{21}$ and $\widetilde{u}_{22}$ are evaluated at a point in between $(x, h(x))$ and (k, k). Since $h(x) < x < k$, we have $V'(k) \leq V'(h(x))$, and so

$$(-\widetilde{u}_{22})(k - h(x)) \leq \widetilde{u}_{21}(k - x)$$

The Hessian of u is negative definite on $N_2(\overline{x})^2$, and so $(-\widetilde{u}_{22}) > 0$, and $[k - h(x)] > [k - x] > 0$, so we get

$$(-\widetilde{u}_{22})(k - x) \leq \widetilde{u}_{21}(k - x)$$

Thus, $\widetilde{u}_{21} \geq -\widetilde{u}_{22}$, and letting $x \to k$, we obtain $u_{21}(k, k) \geq [-u_{22}(k, k)] > 0$. Using (9.15), we have

$$(1 + \delta)|u_{12}(k, k)| < [-u_{22}(k, k)] + [-u_{11}(k, k)]$$

Also, $u_{11}(k,k) < 0$, $u_{22}(k,k) < 0$, since the Hessian of u is negative definite on $N_2(\overline{x})^2$ and $(k,k) \in N_2(\overline{x})^2$. Thus, condition C1 of Section 7 is satisfied, so that by Proposition 7.3, h is differentiable at k, with $|h'(k)| = |\lambda_1| < 1$. But this contradicts the fact that $h(x) < x$ for all x in $(0,k)$, and establishes our claim and hence (i). The claim (ii) can be established similarly.

Thereom 9.1: *Let (Ω, u) be given satisfying condition C2. For every discount factor δ satisfying $\rho_3 \leq \delta < 1$, the optimal policy function, h_δ, and non-zero SOS, k_δ, satisfy*

$$h_\delta^t(x) \to k_\delta \, as \; t \to \infty \tag{9.24}$$

for every $x \in X$, with $x > 0$.

Proof: By Proposition 9.1, there are no period-two optimal cycles generated by h_δ. Thus, by Block and Coppel (1992, Proposition 1, p. 121), given any $x \in X$, $h_\delta^t(x)$ converges to one of the fixed points of h_δ.

There are then two cases to consider: (i) $h_\delta(0) \neq 0$, (ii) $h_\delta(0) = 0$.

In case (i), k_δ is the only fixed point of h_δ, and we are done.

In case (ii), we will be done if we can show that $h_\delta^t(x)$ cannot converge to zero for $x \in X$, $x > 0$.

To this end, we establish first that

$$h_\delta(x) > 0 \text{ for } x > 0 \tag{9.25}$$

To prove (9.25), suppose on the contrary there is some $x \in X$ with $x > 0$, such that $h_\delta(x) = 0$. Then, since $h_\delta(0) = 0$, the sequence $(x, 0, 0, 0, ...)$ is the optimal program from x.

By Proposition 9.2, we must have $x > k_\delta$. Since $k_\delta \in N_2(\overline{x})$, $k_\delta > \overline{x} - \varepsilon_2$; and since $(\overline{x} - \varepsilon_2, \overline{x}) \in N_2(\overline{x})^2$, $(k_\delta, \overline{x}) \in \Omega$ and so $(x, \overline{x}) \in \Omega$. Thus $(x, \overline{x}, \overline{x}, \overline{x}, ...)$ is a program from x. Further, the discounted sum of utilities on this program is at least $u(\overline{x} - \varepsilon_2, \overline{x}) + \delta[u(\overline{x}, \overline{x})/(1-\delta)]$. The discounted sum of utilities on the program $(x, 0, 0, ...)$ is $u(x,0) + \delta[u(0,0)/(1-\delta)]$, which is at most $u(b,0) + \delta[u(0,0)/(1-\delta)]$. Thus, we must have

$$u(b,0) + \delta[u(0,0)/(1-\delta)] \geq u(\overline{x} - \varepsilon_2, \overline{x}) + \delta[u(\overline{x}, \overline{x})/(1-\delta)$$

But this contradicts (9.18), since $\rho_3 \leq \delta < 1$, and establishes (9.25).

In view of (9.25), if $h_\delta^t(x)$ converges to zero, we must have some $T \geq 1$, such that for $t \geq T$, $0 < h_\delta^t(x) < k_\delta$. Further, since $h_\delta^t(x)$ converges to zero we can find some $\tau > T$, such that $h_\delta^{\tau+1}(x) < h_\delta^\tau(x)$. But this contradicts (i) of Proposition 9.2. Thus $h_\delta^t(x)$ cannot converge to zero for $x \in X$, $x > 0$. We infer then that in case (ii) as well, $h_\delta^t(x)$ must converge to k_δ as $t \to \infty$, for every $x \in X$, with $x > 0$. This establishes the turnpike theorem. Q.E.D.

10. Bibliographic Notes

Section 1.

The history of our subject has not been systematically surveyed since the masterly reviews by Koopmans. In the first essay of his celebrated "Three Essays on the State of Economic Science" (1957), Koopmans evaluated the accomplishments of general equilibrium theory and intertemporal allocation theory both in terms of the economic content and the newer mathematical techniques used. The literature on pure capital accumulation oriented optimal growth models was reviewed in Koopmans (1964).

For the theory of Ramsey-type optimal growth models, and some open problems in this area of research, see Koopmans (1967a).

Section 2.

For multisectoral versions of the reduced-form model, see Gale (1967) and McKenzie (1968) (for the case in which future utilities are not discounted), and Sutherland (1970), Stokey and Lucas (1989) (for the case in which future utilities are discounted). In contrast to most presentations of the reduced form model, where the utility function is assumed to be continuous, we develop our model with the assumption that the utility function be bounded and upper semicontinuous. This is because the assumption of continuity on the felicity function in the primitive form of the model (where felicity is derived from consumption) does not always ensure the continuity of the reduced form utility function. This issue was raised by Peleg (1973); for a complete discussion, see Dutta and Mitra (1989a).

Example 2.1 is based on the familiar model developed in continuous time by Cass (1965), Koopmans (1965) and Samuelson (1965). The discrete-time version was presented by Koopmans (1967a). A multisector generalization of this model is contained in Peleg and Ryder (1972). Example 2.2 is based on the continuous-time models of Uzawa (1964) and Srinivasan (1964). For a discrete-time version, see Benhabib and Nishimura (1985) and Boldrin and Montrucchio (1986b). Since then, this has become one of the most familiar models giving rise to cyclical and chaotic optimal behavior.

Example 2.3 has been analyzed in detail in Majumdar and Mitra (1994a) for cyclical and chaotic optimal behavior. A similar model (in continuous time) was examined by Kurz (1968) to study optimal growth with wealth effects. In the renewable resources literature, variations of this model in which the production function exhibits an initial phase of increasing returns followed by decreasing returns (an "S-shape") have been analyzed by Clark (1971), and Majumdar and Mitra (1983).

Example 2.4 is a generalization of the Weitzman example, reported in Samuelson (1973). Variations of this example have been analyzed by Scheinkman (1976), McKenzie (1983), Benhabib and Nishimura (1985) and Mitra and Nishimura (1998), among others.

For variations of Example 2.5, see Wan (1993). The point-input flow-output version of the forestry model (an "orchards" model) was analyzed in Mitra, Ray and Roy (1991). A vintage-capital model, with a similar structure, was examined by Benhabib and Rustichini (1993).

Section 3.

The method of dynamic programming, including the optimality principle, was introduced by Bellman (1957). The rigorous foundation of stochastic dynamic programming was developed by Blackwell (1965), Strauch (1966), Denardo (1967) and Maitra (1968). For generalizations of the results presented in this section to non-stochastic models with a multidimensional state space, see for example Dutta and Mitra (1989b), Stokey and Lucas (1989) and Sorger (1992). One can use a version of the maximum theorem of Berge (1963) to establish that the value function is upper semicontinuous ; we provide a self-contained exposition. Continuity of the value function then follows from the "convex structure" of the model.

Section 4.

For a variety of multisector versions of the basic price characterization results, see Peleg (1970), Peleg and Ryder (1972), Weitzman (1973), Cass and Majumdar (1979), and McKenzie (1986), among others. When there is a (modified) golden-rule (see Section 6), it turns out that the transversality condition (which characterizes the optimality of competitive programs) can be replaced by (an infinite number of) period-by-period conditions. For a discussion of this issue, see Brock and Majumdar (1988) for the undiscounted case, and Dasgupta and Mitra (1988) for the discounted case. These papers, together with other papers dealing with the issue of intertemporal decentralization are contained in Majumdar (1992).

Section 5.

The results, on the continuity of the value and policy functions with respect to changes in the parameters of the model, are closely related to the literature on the maximum theorem, mentioned above in the notes on Section 3. For a survey of results on the general problem of "parametric continuity", see Bank et.al. (1983). The use of supermodularity to obtain monotonicity results on the solution to optimization problems (due to Topkis (1978)) has been extended to the ordinal utility setting by Milgrom and Shannon (1994).

Section 6.

The existence of a stationary optimal stock by dynamic programming methods, in the multisector reduced form model, was established by Sutherland (1970). However, Peleg and Ryder (1974) noted that this stationary optimal stock could be trivial. They proved the existence of a non-trivial stationary optimal stock in a multisectoral model, in which felicity is derived from consumption alone, by establishing the existence of a modified golden-rule. Flynn (1980) and McKenzie (1986) obtained the result in the multisectoral reduced-form model by establishing the existence of a discounted golden-rule stock. Khan and Mitra (1986) established the result using the upper semicontinuity (rather than continuity) of the utility function. Then the result in the Peleg-Ryder (1974) framework can be viewed as a special case, as shown in Dasgupta and Mitra (1990).

Section 7.

The differentiability of the value function, in a stochastic version of the one-sector model(discussed in Example 2.1), was established by Mirman and Zilcha (1975). The result of Benveniste and Scheinkman (1979) was developed for a multidimensional state space setting. For an exposition of the result, see Stokey and Lucas (1989). The technique, developed by Benveniste and Scheinkman, was used by Debreu (1976) in his study of least concave utility functions.

The problem of differentiability of the policy function has been investigated by Boldrin and Montrucchio (1989), Araujo (1991), Santos (1991) and Montrucchio (1998). Araujo addressed the problem in the one-dimensional case. Santos provided the definitive solution to the problem in the multisectoral reduced-form model. Our result applies only to the one-dimensional case and only to the stationary optimal stock. However, we do not assume that the Hessian of the utility function is negative definite, an assumption which is crucial to the methods used in all the above cited papers.

Section 8.

In a continuous-time framework of optimal growth with wealth effects, Kurz (1968) provided an example of non-unique steady-states. Sutherland (1970) provided an example of non-unique stationary optimal stocks directly on a reduced-form model, without discussing any primitive form from which it was derived. In the context of the standard forestry model, Mitra and Wan (1985) provided an example of multiple stationary forests. The uniqueness result for high discount factors, discussed in Section 8.3, is based on Benhabib and Nishimura (1979) and McKenzie (1986). The last part of the proof would have to make use of the Poincare-Hopf index theorem, and the notion of degree of a mapping, in the multidimensional case. In our one-dimensional framework, we can provide a self-contained elementary treatment.

Section 9.

Another example, in which there are optimal period-two cycles for all discount factors close to unity, was developed in the context of Example 2.5 by Wan (1989, 1993). In contrast to the example due to Weitzman, the optimal cycles in his example are boundary cycles, and they are robust to perturbations in the utility function. The earlier counterexamples to the turnpike property, provided by Kurz (1968) and Sutherland (1970), were for discount factors, which were not close to unity. The turnpike property for high discount factors presented in Section 9.2, is crucially dependent on the demonstration (in Proposition 9.1) that there cannot be any period-two optimal cycles. This is really a uniqueness argument, since one is establishing the uniqueness of the solution to a pair of Ramsey-Euler equations. But, unlike the uniqueness problem of Section 8, this problem is two-dimensional, prompting us to use index theory. This theory can be found in Milnor (1965), using the method of differential topology, and in Ortega and Rheinboldt (1970), using purely analytic methods.