Intertemporal Complementarity and Optimality: A Study of a Two-Dimensional Dynamical System*

Tapan Mitra† and Kazuo Nishimura‡

Abstract

We study the underlying structure of the two-dimensional dynamical system generated by a class of dynamic optimization models, which allow for intertemporal complementarity between adjacent periods, but which preserve the time additively separable framework of Ramsey models. Specifically, we identify conditions under which the results of the traditional Ramsey type theory are preserved even when the intertemporal independence assumption is relaxed. Local analysis of this theme has been presented by Samuelson (1971). We establish global convergence results and relate them to the local analysis, by using the mathematical theory of two-dimensional dynamical systems. We also relate the local stability property of the stationary optimal stock to the differentiability of the optimal policy function near the stationary optimal stock, by using the Stable Manifold Theorem.

Journal of Economic Literature Classification Numbers: C61, D90, O41.

Key words: Intertemporal complementarity, supermodularity, single-crossing property, monotonicity, turnpike property, differentiability of policy function, stable manifold theorem.

---

*Earlier versions of this paper were presented in a conference on nonlinear dynamics in Paris in June, 1999, and in a seminar in Kyoto University in January 2001. We are grateful to Jess Benhabib, Cuong Le Van, Roger Farmer and Takashi Kamihigashi for comments received during these presentations. In particular, the final subsection owes much to an observation made by Takashi Kamihigashi. The current version has benefited from comments by five referees of this journal.

†Department of Economics, Cornell University, Ithaca, NY 14853, Email: tm19@cornell.edu.

‡Institute of Economic Research, Kyoto University, Yoshidahonmachi, Sakyoku, Kyoto 606,Japan;E-Mail:nishimura@kier.kyoto-u.ac.jp.

1 Introduction

The theory of optimal intertemporal allocation has been developed primarily for the case in which the objective function of the planner or representative agent can be written as:

$$U(c_1,\, c_2 ...) \equiv \sum_{t=1}^{\infty} \delta^{t-1} w(c_t) \tag{1.1}$$

where c_t represents consumption at date t, w the period felicity function, and $\delta \in (0, 1)$ a discount factor, representing the time preference of the agent.

An objective function like (1.1) leads naturally to the study of dynamic optimization problems of the following "reduced form":

$$\left.\begin{array}{c} Maximize \sum_{t=0}^{\infty} \delta^t u(x_t,\, x_{t+1}) \\ subject\, to\, (x_t,\, x_{t+1}) \in \Omega \;\; for\; t \in \{0, 1, 2, ...\} \\ x_0 = x \end{array}\right\} \tag{1.2}$$

where $\delta \in (0, 1)$ is the discount factor, X is a compact set (representing the state space), $\Omega \subset X \times X$ is a transition possibility set, $u : \Omega \to \mathbb{R}$ is a utility function, and $x \in X$ is the initial state of the system.

The restrictive form of the objective function (1.1) has often been criticized, and alternative forms have been suggested. Since imposing no structure on $U(c_1,\, c_2, ...)$ will yield very little useful information about the nature of optimal programs, the alternative formulations involve some restrictions, of course, and these basically take one of two forms.

First, one can dispense with the *time-additively separable* nature of (1.1), by following Koopmans (1960) and Koopmans, Diamond and Williamson (1964), and postulate that there is an *aggregator function*, A, such that

$$U(c_1,\, c_2, ...) = A(c_1,\, U(c_2,\, c_3, ...)) \tag{1.3}$$

A nice feature of (1.3) is that it preserves the *recursive* nature of the problem inherent in Ramsey-type problems based on (1.1). The restriction is that the *independence of tastes* between periods that was present in (1.1) is also implicit in (1.3). Optimal growth problems with (1.3) as the objective function have been investigated quite extensively, starting with Iwai (1972); a useful reference for this literature is Becker and Boyd (1997).

Second, one can preserve the time-additive separable form, but explicitly model the intertemporal *dependence of tastes* by postulating that the felicity derived by the agent in period t depends on consumption in period $t(c_t)$, but the felicity function itself is (endogenously) determined by past consumption (c_{t-1}). [The fact that "past consumption" is reflected completely in c_{t-1} is a mathematical simplification; consumption in several previous periods can clearly be allowed for at the expense of cumbersome notation and significantly more tedious algebraic manipulations].

This formulation leads to the objective function[1]:

$$U(c_1,\, c_2, ...) = \sum_{t=1}^{\infty} \delta^{t-1} w(c_t, c_{t+1}) \tag{1.4}$$

Models of optimal growth with intertemporal dependence in tastes, in which the objective function is similar to (1.4), have been examined by several authors.[2] To the best of our knowledge, the specific form (1.4) was first used by Samuelson (1971), to capture the essential features of such intertemporal dependence of tastes.

An objective function like (1.4) leads to the study of dynamic optimization problems of the following "reduced form":

$$\left.\begin{array}{c} Maximize \;\; \sum_{t=0}^{\infty} \delta^t u(x_t,\, x_{t+1},\, x_{t+2}) \\ Subject\, to\; (x_t,\, x_{t+1},\, x_{t+2}) \in \Lambda \;\; for\; t \in \{0, 1, 2, ...\} \\ (x_0,\, x_1) = (x, y) \end{array}\right\} \tag{1.5}$$

where $\delta \in (0, 1)$ is the discount factor, X is a compact set, $\Omega \subset X \times X$ is a transition possibility set, $\Lambda = \{(x, y, z) : (x, y) \,\epsilon\, \Omega \text{ and } (y, z) \,\epsilon\, \Omega\}$, $u : \Lambda \to \mathbb{R}$ is a utility function, and $(x, y) \,\epsilon\, \Omega$ is the initial state of the system.

Notice that even under intertemporal dependence in tastes, we have a recursive structure in the dynamic optimization problem (1.5) very much like in (1.2) [and in optimization problems involving (1.3) as the objective function]. The difference is that in dealing with a one capital good model (like the standard one or two-sector models of neoclassical growth theory), the state space is X in problem (1.2), while it is a subset of X^2 in problem (1.5). Thus, for problem (1.2), (optimal) value and policy functions are defined on X, and for problem (1.5), these functions are defined on $\Omega \subset X^2$. In terms of examining the dynamic behavior of optimal programs, we are therefore dealing with a one-dimensional dynamical system for problem (1.2) and a two-dimensional dynamical system for problem (1.5).

The structure of recursive problems like (1.5) are not as well understood as that of (1.2), and we feel that it is worthy of a systematic study. Specifically, one might explore two themes: (i) identifying the conditions under which the results of the traditional Ramsey-type theory are preserved even when the intertemporal independence assumption is relaxed; (ii) examining alternative scenarios in which the asymptotic behavior of an optimal program is qualitatively different (from its traditional Ramsey counterpart) because of the presence of intertemporal complementarity. Local analysis of the first theme has been presented by Samuelson (1971), and of the second by Boyer (1978), and others. Our principal interest in this paper is in establishing global results on the first theme, and in relating them to the local results, by using

[1] This objective function also arises in a somewhat different class of models, which study economic growth with *altruistic preferences*. For this literature, see, for example, Dasgupta (1974), Kohlberg (1976), Lane and Mitra (1981), and Bernheim and Ray (1987). The focus of this literature is however not on the socially optimal solution, but the intergenerational Nash equilibrium solutions.

[2] The earlier literature on this topic includes, among others, Chakravarty and Manne (1968), and Wan (1970). Heal and Ryder (1973) present a continuous-time model which accommodates a more general dependence structure.

the mathematical theory of two-dimensional dynamical systems.[3]

The plan of the paper is as follows. After describing the model in Section 2, we develop the basic properties of the (optimal) value function, V, and the (optimal) policy function, h, in Section 3. A useful tool for our study is the *ϕ-policy function*, defined on X, by

$$\phi(x) = h(x, x) \qquad for \;\; x \,\epsilon\, X \tag{1.6}$$

It is introduced in Section 3, and the circumstances under which it satisfies a "single-crossing condition" are examined.

Section 4 might be considered as providing the global analytical counterpart to Samuelson's (1971) local analysis of "turnpike behavior" in this model. We show that when the (reduced-form) utility function, u, is *supermodular* on its domain, Λ, then the optimal policy function is monotone increasing in both arguments. This property, together with the "single-crossing condition" on ϕ allows us to establish global asymptotic stability of optimal programs with respect to the (unique) stationary optimal stock, by using an interesting stability result for second-order difference schemes.

In Section 5, we provide an analysis of the local dynamics of optimal solutions. To this end, we study the fourth order difference equation, which represents the linearized version of the Ramsey-Euler equations near the stationary optimal stock. This equation yields four characteristic roots and we show how two of them are selected by the optimal solution (assuming that the optimal policy function is continuously differentiable in a neighborhood of the stationary optimal stock). The roots selected by the optimal solution provide information about the speed of convergence of non-stationary optimal trajectories to the stationary optimal stock.

The theory linking the derivative of the optimal policy function to the "dominated" characteristic root associated with the Ramsey-Euler equation, for the optimization problem (1.2) is, of course, well-known. To our knowledge, the corresponding theory for problem (1.5) has not been developed in the literature.

In subsection 5.3, the optimal policy function is shown to be continuously differentiable in a neighborhood of the stationary optimal stock, by using the Stable Manifold Theorem.[4] This validates the conclusions which are reached in Sections 5.1 and 5.2, by assuming this property.

2 Preliminaries

2.1 The Model

Our framework is specified by a *transition possibility set*, Ω, a (reduced form) *utility function*, u, and a *discount factor*, δ. We describe each of these objects in turn.

A *state space* (underlying the transition possibilities) is specified as an interval $X \equiv [0, B]$, where $0 < B < \infty$. The transition possibility set, Ω, is a subset of X^2, satisfying

[3]The second theme is explored in detail in Mitra and Nishimura (2001).

[4]The global differentiability of the optimal policy function for problem (1.2) has been studied by Araujo (1991), Santos (1991) and Montrucchio (1998). The relation of the characteristic roots associated with the optimal policy function to those associated with the Ramsey-Euler equation at the steady state has been studied for problem (1.2) by Araujo and Scheinkman (1977) and Santos (1991).

(A.1) $(0,0)$ and (B,B) are in Ω; if $(0,y) \in \Omega$ then $y=0$.
(A.2) Ω is closed and convex.
(A.3) If $(x,y) \in \Omega$ and $x \leq x' \leq B$, $0 \leq y' \leq y$, then $(x',y') \in \Omega$.
(A.4) There is $(\bar{x},\bar{y}) \in \Omega$ with $\bar{y} > \bar{x}$.

These assumptions are standard in the literature. Note that (A.3) means that the transition possibility set Ω allows *free-disposal*, so long as the stock level does not exceed B. Assumption (A.4) implies the existence of *expansible* stocks.

Notice that for all $x \in [0,B]$, we have $(x,x) \in \Omega$. Associated with Ω is the correspondence $\Psi : X \to X$, given by $\Psi(x) = \{y : (x,y) \,\epsilon\, \Omega\}$. Define the set:

$$\Lambda = \{(x,y,z) : (x,y) \,\epsilon\, \Omega \textit{ and } (y,z) \,\epsilon\, \Omega\}$$

The utility function, u, is a map from Λ to $\mathbb{R}$. It is assumed to satisfy:
(A.5) u is continuous and concave on Λ, and strictly concave in the third argument.
(A.6) u is non-decreasing in the first argument, and non-increasing in the third argument.

In what follows, we will normalize $u(0,0,0)=0$; also, we will denote $\max\limits_{(x,y,z)\epsilon\Lambda} |u(x,y,z)|$ by $\bar{B}$.

The discount factor, δ, reflects how future utilities are evaluated compared to current ones. We assume:
(A.7) $0 < \delta < 1$

2.2 Programs

The *initial condition* (which should be considered to be historically given) is specified by a *pair* (x,y) in Ω. A *program* (x_t) from (x,y) is a sequence satisfying

$$x_0 = x,\ x_1 = y,\ (x_t,\ x_{t+1}) \in \Omega \ \ for\ t \geq 1 \tag{2.1}$$

Thus, in specifying a program, the period 0 and period 1 states are historically given. Choice of future states starts from $t=2$. Notice that for a program (x_t) from $(x,y) \in \Omega$, we have $(x_t,\ x_{t+1},\ x_{t+2}) \in \Lambda$ for $t \geq 0$.

An *optimal program* $(\bar{x}_t)$ from $(x,y) \in \Omega$ is a program from (x,y) satisfying

$$\sum_{t=0}^{\infty} \delta^t u(x_t,\ x_{t+1},\ x_{t+2}) \leq \sum_{t=0}^{\infty} \delta^t u(\bar{x}_t,\ \bar{x}_{t+1},\ \bar{x}_{t+2}) \tag{2.2}$$

for every program (x_t) from (x,y).

Under our assumptions, a standard argument suffices to ensure the existence of an optimal program from every initial condition $(x,y) \in \Omega$. Using Assumptions (A.2) and (A.5), it can also be shown that this optimal program is unique.

2.3 Value and Policy Functions

We can define a *value function*, $V : \Omega \to \mathbb{R}$ by

$$V(x,y) = \sum_{t=0}^{\infty} \delta^t u(\bar{x}_t,\, \bar{x}_{t+1},\, \bar{x}_{t+2}) \tag{2.3}$$

where $(\bar{x}_t)$ is the optimal program from (x, y). Then, V is concave and continuous on Ω.

It can be shown that for each $(x, y) \in \Omega$, the Bellman equation

$$V(x,y) = \max_{(y,z)\epsilon\Omega} [u(x,y,z) + \delta V(y,z)] \tag{2.4}$$

holds. Also, V is the unique continuous function on Ω, which solves the functional equation (2.4).

For each $(x, y) \in \Omega$, we denote by $h(x, y)$ the value of z which maximizes $[u(x, y, z)+\delta V(y, z)]$ among all z satisfying $(y, z) \in \Omega$. Then, a program (x_t) from $(x, y) \in \Omega$ is an optimal program from (x, y) if and only if:

$$V(x_t,\, x_{t+1}) = u(x_t,\, x_{t+1},\, x_{t+2}) + \delta V(x_{t+1},\, x_{t+2}) \qquad for\ t \geq 0 \tag{2.5}$$

This, in turn, holds if and only if

$$x_{t+2} = h(x_t,\, x_{t+1}) \qquad for\ t \geq 0 \tag{2.6}$$

We will call h the (optimal) *policy function*. It can be shown by using standard arguments that h is continuous on Ω.

2.4 Two Examples

2.4.1 Optimal Growth with Intertemporally Dependent Preferences

The example (which follows Samuelson (1971) and Boyer (1978) closely) captures the feature that tastes between periods are intertemporally dependent. Such a model can be described in terms of a *production function*, f, a *welfare function*, w, and a *discount factor*, δ.

Let $X = [0, B]$ be the state space with $0 < B < \infty$. The production function, f, is a function from X to itself which satisfies:

(f) $f(0) = 0, f(B) = B; f$ is increasing, concave and continuous on X.

The welfare function, w, is a function from X^2 to $\mathbb{R}$, which satisfies:

(w) w is continuous and concave on X^2, and strictly concave in the second argument; it is non-decreasing in both arguments.[5]

The discount factor, δ, is as usual assumed to satisfy:

(d) $0 < \delta < 1$.

[5] Boyer (1978) assumes that w is increasing in both arguments. Samuelson (1971) does not; he assumes instead that $w(c, c)$ is increasing in c. It is this latter assumption that is crucial in proving the uniqueness of a stationary optimal stock in this model, and therefore of our "single-crossing property"; see Section 3.3.

A program, in this framework, is described by a sequence $(k_t,\, c_t)$, where k_t denotes the capital stock and c_t the consumption in period t. The initial condition is specified by $(k, c) \geq 0$, where $k + c \leq B$.

Formally, a program $(k_t,\, c_t)$ from (k, c) is a sequence satisfying

$$\left.\begin{array}{lr} (k_1,\, c_1) = (k, c),\; k_{t+1} = f(k_t) - c_{t+1} & for\ t \geq 1 \\ 0 \leq c_{t+1} \leq f(k_t) & for\ t \geq 1 \end{array}\right\} \tag{2.7}$$

An optimal program from (k, c) is a program $(\bar{k}_t,\, \bar{c}_t)$ satisfying

$$\sum_{t=1}^{\infty} \delta^{t-1} w(c_t,\, c_{t+1}) \leq \sum_{t=1}^{\infty} \delta^{t-1} w(\bar{c}_t,\, \bar{c}_{t+1}) \tag{2.8}$$

for every program $(k_t,\, c_t)$ from (k, c).

To reduce the optimality exercise in (2.8) subject to (2.7) to the one in (1.5), we can proceed as follows. First, the transition possibility set, Ω, can be defined as:

$$\Omega = \{(x, y) : x \in X,\, 0 \leq y \leq f(x)\}$$

Second, the reduced form utility function can be defined, for (x, y, z) in Λ as:

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

Finally, the initial condition (k, c) in the example, translates to the initial condition in the framework of Section 2.2 as $(x, y) = (f^{-1}(k_1 + c_1),\, k_1)$. That is, x is the capital stock (in period 0) which produced the output $(k_1 + c_1)$ in period 1, that was split up between consumption (c_1) and capital stock (k_1) in period 1; y is the capital stock in period 1. The choice of consumption decisions, c_t, starts from $t \geq 2$; correspondingly, the state variable, x_t, is determined for $t \geq 2$ by the following equation:

$$x_{t+1} = k_{t+1} = f(k_t) - c_{t+1} \qquad for\ t \geq 1 \tag{2.9}$$

It is worth noting that in the model of Samuelson (1971) there is no maximization with respect to c_1. This is because in order for his problem to be well-posed, one needs to know both k_0 and c_1 (and therefore both x_0 and x_1 in terms of the problem stated in (1.5)) The information about c_1 is needed to *define* the welfare function $w(c_1; c_2)$. That is the welfare in period 2 depends on the choice of c_2, but the welfare function itself is determined endogenously by past consumption (c_1).[6]

[6] Of course, variations of problem (1.5) can arise, where the last line of (1.5) would simply say $x_0 = x$. Solving such a problem would involve solving (1.5) and in addition solving for the "correct" x_1. Clearly, an optimal solution of such a problem must solve (1.5), and therefore inherit all the dynamic properties of such a solution, as described in this paper.

2.4.2 Optimal Harvesting of a Renewable Resource with Delayed Recruitment

The theory of management of renewable resources deals with the issue of optimal harvesting of biological populations, such as various species of marine life. For many species, recruitment to the breeding population takes place only after a delay. Clark (1976) has modeled this phenomenon by describing the population dynamics by a delay-difference equation, instead of the standard first-order difference equation which is commonly used in the literature on renewable resources. We describe a simple version of his model where the delay involved is two periods.[7]

The model can be described formally in terms of a *recruitment function*, F, a *return function*, W, a *survival coefficient*, λ, and a *discount factor*, δ.

The recruitment function, F, is a function from $\mathbb{R}_+$ to itself which satisfies:

(F) $F(0) = 0$; F is increasing, concave and continuous on X; $\lim_{x\to 0}[F(x)/x] > 1, \lim_{x\to\infty}[F(x)/x] = 0$.

The return function, w, is a function from $\mathbb{R}_+$ to $\mathbb{R}$, which satisfies:

(W) W is continuous, non-decreasing and strictly concave on $\mathbb{R}_+$.

The survival coefficient, λ, satisfies:

(s) $0 < \lambda < 1$.

The discount factor, δ, is as usual assumed to satisfy:

(d) $0 < \delta < 1$.

Given (F), there is a unique positive number B, such that $[F(B)/B] = (1-\lambda)$. Then, defining $f(x) = F(x) + \lambda x$ for all $x \in \mathbb{R}_+$, we see that (i) $f(B) = B$, (ii) $B > f(x) > x$ for $x \in (0, B)$, and (iii) $B < f(x) < x$ for $x > B$. Thus, it is natural to choose the state space to be $X = [0, B]$.

A program, in this framework, is described by a sequence $(k_t,\, c_t)$, where k_t denotes the biomass of the adult breeding population and c_t the harvest of this population in period t. The initial condition is specified by $(k, k') \geq 0$, where $k \leq B$ and $k' \leq B$.

Formally, a program $(k_t,\, c_t)$ from (k, k') is a sequence satisfying

$$\left.\begin{array}{rl} (k_0,\, k_1) = (k, k'),\; k_{t+1} = \lambda k_t + F(k_{t-1}) - c_{t+1} & \quad for\ t \geq 1 \\ 0 \leq c_{t+1} \leq \lambda k_t + F(k_{t-1}) & \quad for\ t \geq 1 \end{array}\right\} \tag{2.10}$$

Note that for a program $(k_t,\, c_t)$ from $(k, k') \leq (B, B)$, we have $(k_t,\, c_t) \leq (B, B)$ for all $t \geq 2$, and this justifies our choice of the state space as $X = [0, B]$.

An optimal program from (k, k') is a program $(\bar{k}_t,\, \bar{c}_t)$ satisfying

$$\sum_{t=1}^{\infty} \delta^{t-1} W(c_{t+1}) \leq \sum_{t=1}^{\infty} \delta^{t-1} W(\bar{c}_{t+1}) \tag{2.11}$$

for every program $(k_t,\, c_t)$ from (k, k').

To explain the population dynamics, the adult breeding population k_0 at time 0 yields a "recruitment" to the population in period 2 (that is, after a delay of two periods) of $F(k_0)$. Part of the adult breeding population k_1 at time 1 does not survive beyond period 1; the remaining part is λk_1. The total available output of the renewable resource at time 2 is therefore $F(k_0) + \lambda k_1$. A

[7] The modeling of the recruitment delay as two periods in our formulation of the model of renewable resource management is a mathematical simplification; recruitment delays of longer duration can clearly be allowed for. The corresponding theory is somewhat harder to present and analyze.

part of this resource (c_2) is harvested in period 2. The remainder of the resource ($F(k_0)+\lambda k_1-c_2$) becomes the adult breeding population k_2 at time 2. This process is then repeated indefinitely.

To reduce the optimality exercise in (2.11) subject to (2.10) to the one in (1.5), we can proceed as follows. First, the set, Λ, can be defined as:

$$\Lambda = \{(x,y,z) : x \in X,\ y \in X, 0 \leq z \leq \lambda y + F(x)\}$$

Second, the reduced form utility function can be defined, for (x, y, z) in Λ as:

$$u(x,y,z) = w(\lambda y + F(x) - z)$$

Finally, the initial condition (k, k') in the example, translates to the initial condition in the framework of Section 2.2 as $(x,y) = (k,\ k')$. The choice of consumption decisions, c_t, starts from $t \geq 2$; correspondingly, the state variable, x_t, is determined for $t \geq 2$ by the following equation:

$$x_{t+1} = k_{t+1} = \lambda k_t + F(k_{t-1}) - c_{t+1} \qquad for\ t \geq 1 \tag{2.12}$$

Note that the dynamic optimization problem of the form (1.5) arises in the renewable resource example from the (biological) production side of the model rather than the preference side.

3 Basic Properties of Value and Policy Functions

In this section, we examine some basic properties of the value and policy functions. These properties will be useful in conducting the analysis in the following sections.

3.1 Value Function

We proceed under the following additional assumption:

(A.8) There is $\hat{x}$ in $(0, B)$, such that $(\hat{x},\ \hat{x}/\delta,\ \hat{x}/\delta^2) \in \Lambda$, and $\theta \equiv u(\hat{x},\ \hat{x}/\delta,\ \hat{x}/\delta^2) > u(0,0,0) = 0$.

Assumption (A.8) is a *δ-productivity assumption* jointly on (Λ, u, δ). It is analogous to the δ-productivity assumption in the usual reduced-form model, where it is used to establish the existence of a non-trivial stationary optimal stock.

Lemma 1 *Let $N \geq 2$ be a given positive integer. Defining $x = \delta^N \hat{x}$, we have $(x,\ x/\delta) \in \Omega$, and*

$$V(x,\ x/\delta) \geq [(N-1)\theta/\hat{x}]x \tag{3.1}$$

Proof. Since $(\hat{x},\ \hat{x}/\delta) \in \Omega$ and $(0,0) \in \Omega$, we have $(\delta^n \hat{x},\ \delta^n(\hat{x}/\delta)) \in \Omega$ for $n \geq 1$. Using this observation, the sequence $(x_t) = (x,\ (x/\delta),\ (x/\delta^2), ..., (x/\delta^N),\ (x/\delta^{N+1}), 0, 0, ...)$ is a program from $(x,\ (x/\delta))$. Note that $(x/\delta^N) = \hat{x}$, $(x/\delta^{N+1}) = (\hat{x}/\delta)$, and since $(\hat{x},\ \hat{x}/\delta,\ \hat{x}/\delta^2) \in \Lambda$ by (A.8), we have $(\hat{x},\ \hat{x}/\delta,\ 0) \in \Lambda$ by (A.3), and $u(\hat{x},\ \hat{x}/\delta,\ 0) \geq u(\hat{x},\ \hat{x}/\delta,\ \hat{x}/\delta^2) > 0$. Also $((\hat{x}/\delta),\ 0) \in \Omega$ and $(0,0) \in \Omega$ imply that $(\hat{x}/\delta,\ 0, 0) \in \Lambda$, and $u(\hat{x}/\delta,\ 0, 0) \geq u(0,0,0)$ [by (A.6)] $= 0$. For $0 \leq t \leq N-2$:

$$\begin{aligned} u(x_t,\ x_{t+1},\ x_{t+2}) &= u(x/\delta^t,\ x/\delta^{t+1},\ x/\delta^{t+2}) = u(\hat{x}\delta^{N-t},\ \hat{x}\delta^{N-t-1},\ \hat{x}\delta^{N-t-2}) \\ &\geq \delta^{N-t}u(\hat{x},\ \hat{x}/\delta,\ \hat{x}/\delta^2) + (1-\delta^{N-t})u(0,0,0). \end{aligned}$$

Thus, for $0 \leq t \leq N-2$, $\delta^t u(x_t,\, x_{t+1},\, x_{t+2}) \geq \delta^N u(\hat{x},\, \hat{x}/\delta,\, \hat{x}/\delta^2)$, and we have

$$V(x,\, x/\delta) \geq \sum_{t=0}^{N-2} \delta^t u(x_t,\, x_{t+1},\, x_{t+2}) \geq (N-1)\delta^N \theta = [(N-1)\theta/\hat{x}]x$$

which establishes the Lemma. ■

Proposition 1 *The value function, V, satisfies the property:*

$$[V(x,\, x/\delta)/x] \to \infty \qquad as\ x \to 0 \tag{3.2}$$

Proof. For $(x,\, x/\delta) \in \Omega$, and $0 < \lambda < 1$, we have $V(\lambda x,\, \lambda x/\delta) \geq \lambda V(x,\, x/\delta) + (1-\lambda)V(0,0) = \lambda V(x,\, x/\delta)$. Using Lemma 1, and defining the sequence $\{x(N)\}$ by: $x(N) = \delta^N \hat{x}$ for $N = 2, 3, ...$, we have $[V(x(N),\, x(N)/\delta)/x(N)] \to \infty$ as $N \to \infty$. Then, (3.2) follows since for $x \in [\delta^{N+1}\hat{x},\, \delta^N \hat{x}]$, $V(x,\, x/\delta)/x \geq [V(\delta^N \hat{x},\, \delta^N(\hat{x}/\delta))/\delta^N \hat{x}]$. ■

Proposition 2 *The value function, V, satisfies the property:*

$$[V(x, x)/x] \to \infty \qquad as\ x \to 0 \tag{3.3}$$

Proof. For $0 < x \leq \hat{x}$, we have $(x,\, x/\delta) \in \Omega$, and $(x, 0) \in \Omega$, so $(\delta x + (1-\delta)x,\, \delta(x/\delta) + (1-\delta).0) \in \Omega$; that is $(x, x) \in \Omega$. By concavity of V, we have

$$\begin{aligned} V(x,x) &= V(\delta x + (1-\delta)x,\, \delta(x/\delta) + (1-\delta).0) \\ &\geq \delta V(x,\, x/\delta) + (1-\delta)V(x,0) \\ &\geq \delta V(x,\, x/\delta) \end{aligned}$$

Thus $[V(x,x)/x] \to \infty$ as $x \to 0$ by Proposition 1. ■

3.2 Policy Function

A useful tool, related to the policy function, is the *ϕ-policy function* defined for $x \in X$ by:

$$\phi(x) = h(x, x) \qquad for \qquad x \in X$$

That is, ϕ gives us the optimal policy when the arguments in h happen to take on identical values.

In the standard reduced-form model, if x_t were constant for two successive periods along an optimal program, the constant value would have to be a stationary optimal stock. Here, given $x_{t-1} = x_t = x$ in $X, \phi(x)$ is not necessarily equal to x; in fact, it will typically be different from x. If $\phi(x) = x$, then x would be a stationary optimal stock in the present framework.

We proceed under the following additional assumption:

(A.9) There is $A > 0$, such that for all (x, y, z), $(x',\, y',\, z')$ in Λ, $|u(x, y, z) - u(x',\, y',\, z')| \leq A\,\|(x, y, z) - (x',\, y',\, z')\|$.

Assumption (A.9) is a bounded-steepness assumption on the utility function, and this is ensured by making u Lipschitz-continuous, with Lipschitz constant A. The norm used in (A.9) is the sum-norm; that is, $\|(x,y,z)\| = |x| + |y| + |z|$ for (x,y,z) in $\mathbb{R}^3$. [In the usual reduced-form model, a condition like (A.9) was introduced by Gale (1967), to establish the existence of shadow-prices, associated with optimal programs].

Proposition 3 *There is $a > 0$ such that for all $x \in (0, a)$, $\phi(x) > x$.*

Proof. Suppose, on the contrary, there is a sequence (x^s), such that $x^s \to 0$ as $s \to \infty$, and $x^s > 0$, $\phi(x^s) \leq x^s$ for all s.

Using Proposition 2, we can find $a_1 > 0$, such that for $x \in (0,\, a_1)$, we have

$$[V(x,x)/x] > 4A/(1-\delta) \tag{3.4}$$

Since $x^s \to 0$, we can find s large enough for which $0 < x^s < a_1$. Pick such an x^s and call it x. Then $x \in (0,\, a_1)$ and $\phi(x) \leq x$. Denote $\phi(x)$ by y, and $h(x,y)$ by z.

Since $y \leq x$, and $(y,z) \in \Omega$, we have $(x,z) \in \Omega$, and $(x,y,z) \in \Lambda$,and:

$$\begin{aligned} V(x,x) &\geq u(x,x,z) + \delta V(x,z) \\ &\geq u(x,x,z) + \delta V(y,z) \\ &= [u(x,x,z) - u(x,y,z)] + \delta V(y,z) + u(x,y,z) \\ &= [u(x,x,z) - u(x,y,z)] + V(x,y) \\ &\geq V(x,y) - Ax \end{aligned}$$

the final inequality following from (A.9). We can now write:

$$\begin{aligned} V(x,x) &= u(x,x,y) + \delta V(x,y) \\ &\leq u(x,x,y) + \delta V(x,x) + \delta Ax \end{aligned}$$

so that:

$$\begin{aligned} V(x,x) &\leq [u(x,x,y)/(1-\delta)] + \delta Ax/(1-\delta) \\ &\leq [A(2x+y) + \delta Ax]/(1-\delta) \end{aligned}$$

by using (A.9) again. Thus, using $y \leq x$,we have $[V(x,x)/x] \leq (3+\delta)A/(1-\delta)$ which contradicts (3.4). ■

We now introduce an additional assumption for our next result.

(A.10) $u(B,B,B) \leq u(0,0,0) = 0$.

Assumption (A.10) is an expression of the fact that "inaction" can produce at least as much utility as "excessive action". In the context of the standard aggregative growth model, considered in Section 2.4.1, it translates to the fact that the consumption level associated with the maximum sustainable stock is 0 and so is the consumption associated with the zero stock. Actually the standard aggregative growth model does not model disutility of effort directly. Typically, maintaining high stocks involves considerable effort, which has disutility, and this will reinforce the circumstances under which (A.10) will hold.

Proposition 4 *The ϕ-policy function satisfies*

$$\phi(B) < B \tag{3.5}$$

Proof. Suppose, on the contrary, that $h(B,B) = B$. Then, we have $V(B,B) = u(B,B,B)/(1-\delta) \leq u(0,0,0)/(1-\delta) = 0$. Define the sequence (x_t) as follows:

$$(x_0, x_1) = (B, B);\ x_t = 0 \ \ for\ t \geq 2$$

Then, (x_t) is a program from (B,B), and we have:

$$\sum_{t=0}^{\infty} \delta^t u(x_t, x_{t+1}, x_{t+2}) = u(B,B,0) + \delta u(B,0,0) \geq 0$$

the inequality following from (A.3) and (A.6). This means that $(B,B,0,0,\ldots)$ and $(B,B,B,B,\ldots)$ are both optimal from (B,B). But this contradicts the fact that u is strictly concave in the third argument. This establishes the result. ■

Proposition 5 *There is some $x^* \epsilon\, (0,B)$, such that x^* is a stationary optimal stock; that is,*

$$h(x^*, x^*) = x^* \tag{3.6}$$

Proof. By Proposition 4, $\phi(B) < B$. By Proposition 3, we can find $x \,\epsilon\, (0,B)$, such that $\phi(x) > x$. By continuity of ϕ, there is $x^* \epsilon\, (0,B)$ such that $\phi(x^*) = x^*$. ■

3.3 A Single Crossing Condition

In the following sections, we will find it useful to assume that the ϕ-policy function (introduced in Section 3.2) has the following "single-crossing property":

$$\left.\begin{array}{c}\text{There is } 0 < x^* < B \text{, such that} \\ \phi(x^*) = x^*; x < \phi(x) \text{ for } 0 < x < x^*; x > \phi(x) \text{ for } x > x^*\end{array}\right\} \tag{SC}$$

Given Propositions 3 and 4, there is a stationary optimal stock in $(0,B)$, and the single-crossing property holds if there is a unique stationary optimal stock in $(0,B)$. Note that if $x \in (0,B)$ is a stationary optimal stock, then $h(x,x) = x$, and so:

$$u_3(x,x,x) + \delta u_2(x,x,x) + \delta^2 u_1(x,x,x) = 0 \tag{3.7}$$

In the example discussed in Section 2.4.1, with w and f both C^1, $w_1 > 0$ and $w_2 > 0$, denoting $f(x) - x$ by c, condition (3.7) is satisfied only if:

$$[w_2(c,c) + \delta w_1(c,c)] = \delta f'(x)[w_2(c,c) + \delta w_1(c,c)]$$

This implies that $\delta f'(x) = 1$. Since f is strictly concave, there can be only one stationary optimal stock in $(0,B)$, and the single-crossing property is verified.

If u is C^2 on Λ, then (3.7) has a unique solution if the function:

$$H(x) \equiv u_3(x,x,x) + \delta u_2(x,x,x) + \delta^2 u_1(x,x,x) \tag{3.8}$$

has a negative derivative, wherever it has a zero. This amounts to the condition:

$$[\delta^2 u_{11} + \delta u_{22} + u_{33}] + (\delta^2+1)u_{13} + \delta(\delta+1)u_{12} + (\delta+1)u_{23} < 0 \tag{3.9}$$

being satisfied at any x [the derivatives being evaluated at (x,x,x)] at which $H(x) = 0$. For $\delta \approx 1$, (3.9) is clearly satisfied if u has a negative-definite Hessian.

4 Turnpike Behavior

In this section, we will provide sufficient conditions under which one can establish global asymptotic stability of the stationary optimal stock (turnpike property). This demonstrates that "one can relax the independence assumption somewhat and still derive the usual known results", a point indicated earlier by Samuelson (1971), using local analysis around the turnpike. A crucial role in our global analysis is played by the assumption of *supermodularity* of the utility function[8] in its *three* variables, a concept we define below analogously to the more familiar two variable case.

4.1 Supermodularity of the Utility Function

A function $G : \Omega \to \mathbb{R}$ is *supermodular* if whenever (x,y), $(x',y') \in \Omega$ with $(x',y') \geq (x,y)$, we have

$$G(x,y) + G(x',y') \geq G(x',y) + G(x,y')$$

provided (x',y) and $(x,y') \in \Omega$. If G is C^2 on Ω, then it is well-known that G is supermodular on Ω if and only if $G_{12} \geq 0$ on Ω

In our case, the utility function, $u : \Lambda \to \mathbb{R}$ is a function of three variables, and we may define supermodularity of it as follows. In the C^2 case, we would now like to have all the three cross-partials of u to be non-negative; that is u_{12}, u_{13} and $u_{23} \geq 0$ on Λ. In the general (not necessarily differentiable) case, this translates to the following definition.

The utility function $u : \Lambda \to \mathbb{R}$ is called *supermodular* on Λ if whenever (x,y,z), $(x',y',z') \in \Lambda$ with $(x',y',z') \geq (x,y,z)$, we have

(i) $u(x,y,z) + u(x',y',z') \geq u(x',y,z) + u(x,y',z')$ provided (x',y,z) and $(x,y',z') \in \Lambda$

(ii) $u(x,y,z) + u(x',y',z') \geq u(x,y',z) + u(x',y,z')$ provided (x,y',z) and $(x',y,z') \in \Lambda$

(iii) $u(x,y,z) + u(x',y',z') \geq u(x,y,z') + u(x',y',z)$ provided (x,y,z') and $(x',y',z) \in \Lambda$

If u is C^2 on Λ with u_{12}, u_{12} and $u_{23} \geq 0$ on Λ, then (i) can be verified as follows: $A = [u(x',y',z') - u(x,y',z')] - [u(x',y,z) - u(x,y,z)] = \int_x^{x'} u_1(t,y',z')dt - \int_x^{x'} u_1(t,y,z)dt$. Now $u_1(t,y',z') - u_1(t,y,z) \geq 0$ for all $t \in [x,x']$, since $u_{12} \geq 0$ and $u_{13} \geq 0$, $(y'-y) \geq 0$ and $(z'-z) \geq 0$. Thus, we get $A \geq 0$, establishing (i). Conditions (ii) and (iii) can be verified similarly.

[8] The supermodularity concept is due to Topkis (1968). A nice exposition of the concept in the two variable case, and its relation to the non-negativity of the cross partial derivative is given in Ross (1983). Benhabib and Nishimura (1985) introduced its use in optimal economic dynamics in the two variable case in the form of this derivative condition. More recent comprehensive studies involving the concept of supermodularity can be found in Amir, Mirman and Perkins (1991), Amir (1996) and Topkis (1998).

Both of the above definitions are, of course, special cases of the general definition of supermodularity of a function on a lattice, as given by Topkis (1968, 1998).

4.2 An Example

In order to understand the restriction imposed by the assumption of supermodularity of u, we consider the assumption in the context of the example discussed in Section 2.4.1, when w and f are both C^2 on their domains.

We can calculate the first-order partial derivatives of u as follows:

$$\begin{aligned} u_1(x,y,z) &= w_1(f(x)-y,\ f(y)-z)f'(x) \\ u_2(x,y,z) &= w_2(f(x)-y,\ f(y)-z)f'(y) - w_1(f(x)-y,\ f(y)-z) \\ u_3(x,y,z) &= -w_2(f(x)-y,\ f(y)-z) \end{aligned}$$

Since $w_1 > 0$ and $w_2 > 0$, it follows that $u_1 > 0$ and $u_3 < 0$, as required in (A.6).

The second-order cross partial derivatives of u can be calculated as follows:

$$\begin{aligned} u_{12}(x,y,z) &= [w_{12}(f(x)-y,\ f(y)-z)f'(y) - w_{11}(f(x)-y,\ f(y)-z)]f'(x) \\ u_{13}(x,y,z) &= -f'(x)w_{12}(f(x)-y,\ f(y)-z) \\ u_{23}(x,y,z) &= w_{12}(f(x)-y,\ f(y)-z) - w_{22}(f(x)-y,\ f(y)-z)f'(y) \end{aligned}$$

Thus, in order for u to be supermodular, (i) we need the marginal utility of present consumption to be declining in past consumption ($w_{12} < 0$), sometimes referred to as Edgeworth-Pigou substitutability, and (ii) we need the magnitude of this cross effect ($-w_{12}$) to be "small" relative to the magnitudes of the own effects ($-w_{11}$) and ($-w_{22}$). We state requirement (ii) loosely, since the magnitude of the marginal product of capital is involved beside the second-order derivatives of w. However, the requirement (ii) can be seen most transparently at the steady state, where $\delta f'(x^*) = 1$. There, the requirement of supermodularity translates to the condition that in the symmetric matrix:

$$W = \begin{bmatrix} \delta^2(-w_{11}) & \delta(-w_{12}) \\ \delta(-w_{12}) & (-w_{22}) \end{bmatrix}$$

the diagonal terms dominate the off-diagonal terms.

We now provide a specific example of the framework discussed in Section 2.4.1, to show that all the assumptions made on the reduced-form model can be verified with suitable restrictions on the parameters of the primitive-form. Consider the production function, f, defined by:

$$f(x) = px - qx^2 \qquad for\ x \in [0,\ (p-1)/q] \equiv [0, B]$$

(where $1 < p < 2$ and $q > 0$), and the welfare function, w, defined by:

$$w(c,d) = ac - bc^2 + \alpha d - \beta d^2 - \theta\, cd \qquad for\ (c,d) \in X^2$$

where $a > 0$, $b > 0$, $\alpha > 0$, $\beta > 0$ and $\theta > 0$.Note that at $x = B = (p-1)/q$, we have $[f(x)/x] = p - qx = p - q[(p-1)/q] = 1$. Also, $f'(x) = p - 2qx$ for all $x \in X$, so $f'(0) = p$ and $f'(B) = p - 2q[(p-1)/q] = 2 - p$. Since $1 < p < 2$, we have $f'(0) > 1 > f'(B) > 0$.

To ensure that w is increasing in both components of consumption, we impose the following restrictions:

$$a - (\theta + 2b)B > 0; \quad \alpha - (\theta + 2\beta)B > 0 \tag{R1}$$

These restrictions ensure that $u_1(x, y, z) > 0$ and $u_3(x, y, z) < 0$ on Λ, since $f'(x) > 0$ on X.

Notice that $w_{11} = -2b < 0$ and $w_{22} = -2\beta < 0$, so, to ensure concavity of w, we can assume

$$\theta^2 < 4b\beta \tag{R2}$$

This ensures that u is concave on Λ, since f is concave on X. Further, since $u_{33}(x, y, z) = w_{22}(f(x) - y,\ f(y) - z) = -2\beta < 0$, u is strictly concave in its third argument.

We have $w_{12} = -\theta < 0$ so that $u_{13} > 0$ on Λ. To ensure that $u_{23} > 0$ on Λ, we assume that $(-w_{22})f'(B) > (-w_{12})$; that is

$$2\beta(2 - p) > \theta \tag{R3}$$

Finally, to ensure that $u_{12} > 0$ on Λ, we assume that $(-w_{11}) > (-w_{12})f'(0)$; that is,

$$2b > \theta p \tag{R4}$$

Thus, under the restrictions (R1) - (R4), assumptions (w), (f) are satisfied, and so are Assumptions (A.1) - (A.6). Further, u is supermodular on Λ.

For specific numerical values of the parameters, ensuring that all the above restrictions are simultaneously satisfied, take $p = (3/2)$, $q = (1/2)$, so that $B = 1$ and $X = [0, 1]$. Choosing $b = \beta = 1$, $a = 3$, $\alpha = 5$, and $\theta = (1/2)$, it is easy to check that the restrictions (R1) - (R4) are satisfied.

4.3 Monotonicity of the Policy Function

The principal result (Theorem 1) of this subsection is that if the utility function is supermodular then the (optimal) policy function is monotone non-decreasing in each component.

In the case usually treated, where the reduced form utility function is a function of two variables, if the utility function is supermodular, then the policy function is monotone non-decreasing, and this can be established by ensuring that the value function (a function of a single variable) is monotone non-decreasing. This property of the value function is straightforward, given the free-disposal property of the transition possibility set and the fact that the utility function is monotone non-decreasing in its first argument.

In the present context, the value function is a function of two variables, and we need to show that the value function is *supermodular* in these two variables (Proposition 6), when the utility function is supermodular in its three variables. To obtain the supermodularity of the value function from the supermodularity of the utility function, the natural route suggested is to establish supermodularity for each finite-horizon value function, and then obtain this property for the infinite horizon value function as a limit of the finite-horizon ones. The first part of this two-step procedure (Lemma 2) follows from the general result of Topkis (1968), so we state the result without a proof.

Lemma 2 *Let $G : \Omega \to \mathbb{R}$ be a concave, continuous and supermodular function on Ω. If u is* *supermodular on Λ, then the function $H : \Omega \to \mathbb{R}$ given by*

$$H(x,y) = \max_{z \epsilon \Psi(y)} [u(x,y,z) + \delta G(y,z)] \tag{P}$$

is well-defined, and is a concave, continuous and supermodular function on Ω.

Proposition 6 *If u is supermodular on Λ, then V is supermodular on Ω.*

Proof. Define a sequence of functions, $V^t : \Omega \to \mathbb{R}$ given by

$$V^0(x,y) = u(x,y,0) \text{ and } V^{t+1}(x,y) = \max_{z \epsilon \Psi(y)} [u(x,y,z) + \delta V^t(y,z)]$$

Then V^0 is a concave, continuous and supermodular function on Ω. Using Lemma 2, V^t is a concave, continuous and supermodular function on Ω for each $t \geq 0$.

Since $|u(x,y,z)| \leq \bar{B}$ on Λ, we have $|V^t(x,y)| \leq \bar{B}/(1-\delta)$ on Ω for all $t \geq 0$. To see this, note that it is clearly true for $t = 0$. Assuming this is true for $t = T \geq 0$, we have

$$\left|V^{T+1}(x,y)\right| \leq \bar{B} + \delta[\bar{B}/(1-\delta)] = \bar{B}/(1-\delta)$$

Thus $|V^t(x,y)| \leq \bar{B}/(1-\delta)$ on Ω for all $t \geq 0$ by induction.

We now proceed to show that $V^{t+1}(x,y) \geq V^t(x,y)$ for $t \geq 0$, for all $(x,y) \in \Omega$. For $t = 0$, we have

$$\begin{aligned} V^1(x,y) &= \max_{z \epsilon \Psi(y)} [u(x,y,z) + \delta V^0(y,z)] \\ &\geq u(x,y,0) + \delta V^0(y,0) \\ &= u(x,y,0) + \delta u(y,0,0) \\ &\geq u(x,y,0) = V^0(x,y) \end{aligned}$$

since u is non-decreasing in its first argument and $u(0,0,0) = 0$.

Suppose $V^{t+1}(x,y) \geq V^t(x,y)$ for $t = 0, ..., T$ where $T \geq 0$. We now show that the inequality must hold for $t = T+1$ as well. Let $\bar{z}$ be the solution of the maximization problem

$$\max_{z \epsilon \Psi(y)} [u(x,y,z) + \delta V^T(y,z)]$$

given $(x,y) \in \Omega$. Then, by definition of V^{T+2}, we have

$$\begin{aligned} V^{T+2}(x,y) &\geq u(x,y,\bar{z}) + \delta V^{T+1}(y,\bar{z}) \\ &\geq u(x,y,\bar{z}) + \delta V^T(y,\bar{z}) \\ &= V^{T+1}(x,y) \end{aligned}$$

This completes the induction proof.

For each $(x,y) \in \Omega$, define

$$\bar{V}(x,y) = \lim_{t \to \infty} V^t(x,y)$$

Then $\bar{V}$ is well-defined and is a concave, continuous and supermodular function on Ω.

Given $(x, y) \in \Omega$, let z^t be the solution to the maximization problem:

$$\max_{z \epsilon \Psi(y)} [u(x, y, z) + \delta V^t(y, z)]$$

Then, we have

$$V^{t+1}(x, y) = u(x, y, z^t) + \delta V^t(y, z^t)$$

The sequence $\{z^t\}$ is bounded, and has a convergent subsequence, converging to some $\bar{z}$; clearly $\bar{z} \in \Psi(y)$. For the subsequence on which z^t converges to $\bar{z}$, taking limits we have

$$\bar{V}(x, y) = u(x, y, \bar{z}) + \delta \bar{V}(y, \bar{z}) \tag{4.12}$$

Also, for all $z \in \Psi(y)$, we have

$$V^{t+1}(x, y) \geq u(x, y, z) + \delta V^t(y, z)$$

and so

$$\bar{V}(x, y) \geq u(x, y, z) + \delta \bar{V}(y, z) \tag{4.13}$$

Using (4.12) and (4.13) we have

$$\bar{V}(x, y) = \max_{z \epsilon \Psi(y)} [u(x, y, z) + \delta \bar{V}(y, z)]$$

Thus, $\bar{V}$ is the value function, V, of problem (2.4), and V is supermodular on Ω. ■

Theorem 1 *If u is supermodular on Λ, then h is non-decreasing in each component.*

Proof. Let (x, y) and $(x', y') \in \Omega$ with $(x', y') \geq (x, y)$. Define $z = h(x, y)$ and $z' = h(x', y')$. We claim that $z' \geq z$. Suppose, on the contrary, that $z' < z$. We know that

$$\begin{aligned} V(x, y) &= u(x, y, z) + \delta V(y, z) \\ V(x', y') &= u(x', y', z') + \delta V(y', z') \end{aligned}$$

Since $(y, z) \,\epsilon\, \Omega$ and $z' < z$, $(y, z') \,\epsilon\, \Omega$, and

$$V(x, y) > u(x, y, z') + \delta V(y, z')$$

Since $(y, z) \,\epsilon\, \Omega$ and $y' \geq y$, $(y', z) \,\epsilon\, \Omega$, and

$$V(x', y') > u(x', y', z) + \delta V(y', z)$$

Thus, we get

$$\begin{aligned} & [u(x, y, z) + u(x', y', z')] + \delta[V(y, z) + V(y', z')] \\ > \; & [u(x, y, z') + u(x', y', z)] + \delta[V(y, z') + V(y', z)] \end{aligned} \tag{4.14}$$

Since u is supermodular on Λ, and $(x', y') \geq (x, y)$ and $z > z'$,

$$u(x, y, z') + u(x', y', z) \geq u(x, y, z) + u(x', y', z') \tag{4.15}$$

Since V is supermodular on Ω, and $y' \geq y$ and $z > z'$,

$$\delta[V(y, z') + V(y', z)] \geq \delta[V(y, z) + V(y', z')] \tag{4.16}$$

Adding (4.15) and (4.16), we contradict (4.14). ■

4.4 Global Dynamics

In this section we study the global dynamics of the two-dimensional dynamical system, (Ω, Γ) where Γ is a map from Ω to Ω given by

$$\Gamma(x, y) = (y, h(x, y))$$

For $(x, y) \in \Omega$, we have $h(x, y) \in \Psi(y)$, and so $(y, h(x, y)) \in \Omega$.

We maintain the assumption that u is supermodular on Λ, and so h is non-decreasing in both its arguments. We also maintain the single-crossing condition on ϕ, introduced in Section 3c.

The principal result of this subsection (Theorem 2) is that if (x_t) is an optimal program from (x, y), where $(x, y) \in \Omega$ and $(x, y) >> 0$, then x_t converges to x^* as $t \to \infty$, thus exhibiting global asymptotic stability ("turnpike property").

Theorem 2 *Let (x_t) be an optimal program from $(x, y) \in \Omega$ with $(x, y) >> 0$. Then $\lim_{t\to\infty} x_t = x^*$.*

Proof. Define $m = \min\{x_0, x_1, x^*\}$ and $M = \max\{x_0, x_1, x^*\}$, where x^* is given by the single-crossing condition (SC).

We show that $x_t \geq m$ for all $t \geq 0$. This is clear for $t = 0, 1$. Suppose $x_t \geq m$ for $t = 0, 1, ...T$, where $T \geq 1$, then

$$x_{T+1} = h(x_{T-1}, x_T) \geq h(m, m) \geq m \tag{4.17}$$

The first inequality in (4.17) follows from the monotonicity of h in both arguments, the definition of m, and the fact that x_{T-1} and x_T are at least as large as m. The second inequality follows from the fact that $m \leq x^*$and condition (SC). This establishes by induction that $x_t \geq m$ for $t \geq 0$.

We show that $x_t \leq M$ for all $t \geq 0$. This being clear for $t = 0, 1$, suppose $x_t \leq M$ for $t = 0, 1, ..., T$, were $T \geq 1$. Then

$$x_{T+1} = h(x_{T-1}, x_T) \leq h(M, M) \leq M \tag{4.18}$$

The first inequality in (4.18) follows from the monotonicity of h in both arguments, the definition of M, and the facts that $x_{T-1} \leq M$, $x_T \leq M$ by hypothesis. The second inequality follows from the fact that $M \geq x^*$ and condition (SC). This establishes by induction that $x_t \leq M$ for $t \geq 0$.

We also note that

$$x_{t+1} = h(x_{t-1}, x_t) \leq h(B, B) = \phi(B) \tag{4.19}$$

Define $a = \liminf_{t\to\infty} x_t$. We claim that $a \geq x^*$. Otherwise, if $a < x^*$, then using $a \geq m > 0$, we have $\phi(a) > a$ and so we can find $\varepsilon > 0$ such that $(a - \varepsilon) > 0$ and $\phi(a - \varepsilon) > a + \varepsilon$. By definition of a, we can find N such that for $t \geq N$, $x_t \geq (a - \varepsilon)$. Thus, for $t \geq N$,

$$x_{t+2} = h(x_t, x_{t+1}) \geq h(a - \varepsilon, a - \varepsilon) = \phi(a - \varepsilon) > a + \varepsilon$$

But this means that $\liminf_{t\to\infty} x_t \geq a + \varepsilon$, a contradiction. Thus, we must have $a \geq x^*$.

Define $A = \limsup_{t\to\infty} x_t^*$. We claim that $A \leq x^*$. Suppose, on the contrary, $A > x^*$. We know that $A \leq \phi(B)$ [by (4.19)]$< B$ [by Proposition 4]. Using condition (SC) we have $\phi(A) < A$, and so we can find $\varepsilon > 0$ such that $(A+\varepsilon) < B$, and $\phi(A+\varepsilon) < (A-\varepsilon)$. By definition of A, we can find N such that for $t \geq N$, $x_t \leq (A+\varepsilon)$. Thus for $t \geq N$,

$$x_{t+2} = h(x_t, x_{t+1}) \leq h(A+\varepsilon, A+\varepsilon) = \phi(A+\varepsilon) < A - \varepsilon$$

But this means that $\limsup_{t\to\infty} x_t \leq A - \varepsilon$, a contradiction. Thus, we must have $A \leq x^*$.

Since $A \geq a$, we have

$$x^* \geq A \geq a \geq x^* \tag{4.20}$$

which proves that $a = A = x^*$, and so (x_t) converges and $\lim_{t\to\infty} x_t = x^*$. ■

Remark 1 *The style of proof is similar to that used in Hautus and Bolis (1979), but since the domain of definition of h and ϕ are different in our framework from theirs, we cannot appeal directly to their result.*

4.5 Remarks on Models of Habit Formation

The literature on habit formation [9] studies optimization problems of the type described in Section 2.4.1. The model of Boyer (1978) on habit formation, where utility is assumed to be increasing both in current and in past consumption, can be treated as a special case of the model we described in Section 2.4.1.[10] However, in many models of habit formation, utility is assumed to be increasing in present consumption, but *decreasing* in past consumption. The idea is that a high consumption in the past means that a person gets used to a higher standard, and therefore this has a negative effect on her evaluation of current consumption.

Assumption (w) in the example described in Section 2.4.1 (and more generally assumption (A.6) of the reduced-form model described in Section 2.1) rules out such environments of habit formation. However, the methods and results of our paper continue to be applicable to *some* of these environments. We elaborate on this remark by presenting an example of habit formation

[9] See Boyer (1978), Abel (1990) and Deaton (1992) and the references cited by them for the main contributions to this literature.

[10] He does not look at the corresponding reduced-form model, and does not assume conditions on the primitive form which would ensure the supermodularity of the reduced-form utility function. Thus, in his model, unlike ours, it is "possible to experience cycles in consumption, investment, capital and the interest rate" (Boyer (1978, p.594)).

where utility is assumed to be increasing in present consumption and decreasing in past consumption, and yet the main monotonicity and global asymptotic stability results of our paper continue to hold.

The framework of the example is similar to the one described in Section 2.4.1, with a *production function*, f, satisfying assumption (f), a *discount factor*, δ, satisfying assumption (d); however, the *welfare function*, w, is a function from X^2 to $\mathbb{R}$, which satisfies:

(w') w is continuous and concave on X^2, and strictly concave in the second argument; it is non-decreasing in the second argument and non-increasing in the first argument.

We now provide a specific example of this framework, which is a variation of the example discussed in Section 4.2..Consider the production function, f, defined by:

$$f(x) = px - qx^2 \qquad for\ x \in [0,\ (p-1)/q] \equiv [0, B] = X$$

Here $p = (3/2)$ and $q = (1/2)$, so that $B = 1$.

The welfare function, w, is defined by:

$$w(c,d) = A[d/(1+d)] - ac^b \qquad for\ (c,d) \in X^2$$

where $a \in (0,1)$, $b \geq 2$, and $A > 4ab$.

The discount factor, δ, is chosen to be in $(0.8, 1)$.

For any $x \in (0,1)$, we have $f(x) > x$, and so the stationary program $(x, x, x, ...)$ is feasible. Denoting $f(x) - x$ by c, we note that $c \in (0,1)$ is the constant consumption along this program. Given the form of w, we have $w(c,c) > 0$.

Note that w is C^2 on X^2; the partial derivatives of w can be calculated as follows:

$$\begin{aligned}
w_1(c,d) &= -abc^{b-1} < 0 \\
w_2(c,d) &= A/(1+d)^2 > 0 \\
w_{11}(c,d) &= -b(b-1)ac^{b-2} < 0 \\
w_{12}(c,d) &= 0 \\
w_{22}(c,d) &= -2A/(1+d)^3 < 0
\end{aligned}$$

Thus, w is clearly increasing in d and decreasing in c on X^2. Further, it is strictly concave in (c,d) on X^2. It can be checked that for all $c \in (0,1)$, we have $w_1(c,c) + w_2(c,c) > 0$, as required in the study of Samuelson (1971) and the habit formation model of Sundaresan (1989).

The properties (A.1)-(A.4) of the corresponding reduced form model (Λ, u, δ) can be verified quite easily. To verify other properties, we can compute the relevant first and second-order partial derivatives of u as in Section 4.2. Since $w_1 < 0$ and $w_2 > 0$, it follows that $u_1 < 0$ and $u_3 < 0$, while $u_2 > 0$. Note that (A.6) is clearly violated. By definition of u, it is clearly concave (given concavity of w and f) and continuous on Λ, so that (A.5) is satisfied.

Since $w_{12} = 0$ and $w_{11} < 0$, $w_{22} < 0$, we have $u_{13} = 0$ and $u_{12} > 0, u_{23} > 0$. Thus, u is supermodular on Λ.[11]

[11] In order to keep our example simple, we have used a form for $w(c,d)$ for which $w_{12} = 0$. However, this is not essential to the example. One can allow for functions $w(c,d)$ in which $w_{12} < 0$, and the cross effect is small relative to the direct effects of w_{11} and w_{22}, (as explained in Section 4.2) and still preserve the main results of

Assumption (A.7) is clearly satisfied by definition of δ. To verify (A.8), define $\hat{x} = (\delta/4)$. Then $(\hat{x}/\delta) = (1/4)$ and $(\hat{x}/\delta^2) = (1/4\delta)$. Note that $f(\hat{x}/\delta) = f(1/4) = (11/32) > (10/32) > (1/4\delta) = (\hat{x}/\delta^2)$; it follows that, $f(\hat{x}) \geq \delta f(\hat{x}/\delta) > (\hat{x}/\delta)$. Thus, $(\hat{x}, (\hat{x}/\delta), (\hat{x}/\delta^2)) \in \Lambda$. Note that for $x \in (0, \hat{x}/\delta)$, defining $g(x) = f(x) - (x/\delta)$, we have $g'(x) = (3/2) - x - (1/\delta) > (3/2) - (1/4) - (5/4) = 0$. Thus, we have $f(\hat{x}/\delta) - (\hat{x}/\delta^2) > f(\hat{x}) - (\hat{x}/\delta)$, and so using $w_2 > 0$, we have:

$$u(\hat{x}, (\hat{x}/\delta), (\hat{x}/\delta^2)) \geq w(f(\hat{x}) - (\hat{x}/\delta), f(\hat{x}) - (\hat{x}/\delta)) > 0$$

the last inequality following from (7). Thus, (A.8) is verified. Assumption (A.9) is clearly satisfied since both f and w have bounded steepness. Finally, Assumption (A.10) is satisfied, since $u(0,0,0) = 0 = u(B,B,B)$. To summarize, in this example, all the assumptions except (A.6) are satisfied; further, u is supermodular on Λ.

Since the analysis in our paper relies at various points on the use of assumption (A.6), our methods are not directly applicable to this example of habit formation. However, slight modifications of our methods can be used to verify that (i) the policy function satisfies the single-crossing property; (ii) the value function is supermodular on its domain, and the policy function is non-decreasing in each of its arguments. Thus, Theorem 2, which uses only these properties of the model, continues to be valid in this example of habit formation. The details of this verification can be found in Mitra and Nishimura (2003).

We should add that there are clearly models of habit formation which *cannot* be analyzed in terms of the methods used in our paper, and the results of our paper *do not* apply to those frameworks. For example, one might follow Abel (1990) and consider a particular specification of the habit-formation model where:

$$w(c_t, c_{t+1}) = (1/(1-\alpha))(c_{t+1}/c_t)^{(1-\alpha)} \text{ where } \alpha \in (0,1)$$

In this case, not only is $w_1 < 0$ and $w_2 > 0$, so that the corresponding reduced-form model violates (A.6), but w itself is *not* a concave function of (c_t, c_{t+1}), so that the corresponding reduced-form model also violates (A.5). The dynamic optimization problem in (1.5) is then one involving a non-concave objective function. This takes one beyond the scope of environments that can be handled with the methods used in our paper.

5 Local Dynamics

In this section, we provide an analysis of the local dynamics of optimal solutions near a stationary optimal stock. To this end, we study (in subsection 5.1) the behavior of the optimal policy function (assuming that it is continuously differentiable in a neighborhood of the stationary optimal stock) and obtain restrictions on the two characteristic roots associated with the linearized version of it near the stationary optimal stock. Next, we show (in subsection 5.2) that each of these characteristic roots must also be a characteristic root of the linearized version of the Ramsey-Euler equation near the stationary optimal stock. We then examine the fourth order difference equation, which represents the linearized version of the Ramsey-Euler equations near the stationary optimal stock, and we show which two of them are selected by the optimal

---

this example.

solution. The roots selected by the optimal solution provide information about the speed of convergence of non-stationary optimal trajectories to the stationary optimal stock. The assumption of supermodularity of the utility function is not used in the above analysis.

In subsection 5.3, the optimal policy function is shown to be continuously differentiable in a neighborhood of the stationary optimal stock, by using the Stable Manifold Theorem. This provides a rigorous basis for the analysis carried out in subsections 5.1 and 5.2.[12]

5.1 Characteristic Roots Associated with the Optimal Policy Function

We proceed with our local analysis of the optimal policy function by making strong smoothness assumptions.

We assume that there is $\varepsilon > 0$ such that the utility function is C^2 in a neighborhood $N \equiv Q^3$ of (x^*, x^*, x^*), [where $Q = (x^* - \varepsilon, x^* + \varepsilon)$] with $u_1 > 0$, $u_3 < 0$ and $u_{13} > 0$, and a negative-definite Hessian on N. Further, we assume that there is a neighborhood M' of (x^*, x^*) on which V is C^2 and h is C^1.[13] Clearly, we can choose a smaller neighborhood M of M' such that for all (x, y) in M, $(x, y, h(x, y))$ is in N and $(y, h(x, y))$ is in M'.

In terms of the example of Section 2.4, the restriction $u_{13} > 0$ is satisfied if w is C^2 with $w_{12} < 0$ [and f is C^1, with $f' > 0$]. This restriction is quite important: it implies that the policy function is monotone increasing in the first argument on M.

Proposition 7 *The policy function, h, satisfies $h_1(x, y) > 0$ for $(x, y) \in M$.*

Proof. Let $(x, y) \in M$. Then $h(x, y)$ solves the maximization problem:

$$\underset{(y,z)\in\Omega}{Max}[u(x, y, z) + \delta V(y, z)]$$

Since $(y, h(x, y)) \in M'$ and $(x, y, h(x, y))$ is in N,

$$u_3(x, y, h(x, y)) + \delta V_2(y, h(x, y)) = 0 \tag{5.1}$$

This is an identity in $(x, y) \in M$, and so, differentiating with respect to x,

$$u_{31}(x, y, h(x, y)) + u_{33}(x, y, h(x, y))h_1(x, y) + \delta V_{22}(y, h(x, y))h_1(x, y) = 0$$

We have $V_{22} \leq 0$ [by concavity of V], and $u_{33} < 0$ [since the Hessian of u is negative definite]; thus $u_{33}(x, y, h(x, y)) + \delta V_{22}(y, h(x, y)) < 0$, and so $h_1(x, y) > 0$. ■

If $x^* > 0$ is the unique positive stationary optimal stock, another useful property of the optimal policy function may be obtained, namely: $h_1(x^*, x^*) + h_2(x^*, x^*) \leq 1$. Recall that the circumstances under which there is a unique positive stationary optimal stock were discussed in connection with the single-crossing condition in Section 3.3.

[12] Our proof of the continuous differentiability of the optimal policy function near the stationary optimal stock involves using the result as well as the method of proof of Theorem 2. Since the latter result was proved by us under the assumption of supermodularity of the utility function, we are not able to totally dispense with the supermodularity assumption in Section 5.

[13] The circumstances under which these smoothness assumptions hold are given in Section 5.3.

Proposition 8 *Suppose x^* is the unique positive stationary optimal stock. Then*

$$h_1(x^*, x^*) + h_2(x^*, x^*) \leq 1$$

Proof. Since $[h(x, x) - x] \geq 0$ for $0 \leq x \leq x^*$, and $[h(x^*, x^*) - x^*] = 0$, we must have $[h(x, x) - x]$ minimized at $x = x^*$ among all $x \in [0, x^*]$. Thus,

$$h_1(x^*, x^*) + h_2(x^*, x^*) - 1 \leq 0$$

which establishes the result. ■

Given the non-linear difference equation

$$x_{t+2} = h(x_t, x_{t+1})$$

the linear difference equation associated with it (near the stationary optimal stock, x^*) is given by

$$a_{t+2} = qa_t + pa_{t+1} \tag{5.2}$$

where q denotes $h_1(x^*, x^*)$ and p denotes $h_2(x^*, x^*)$, and a_t is to be interpreted as $(x_t - x^*)$ for $t \geq 0$.

The characteristic equation associated with the equation (5.2) is

$$\lambda^2 = q + p\lambda \tag{5.3}$$

Denoting by λ_1 and λ_2 the roots of (5.3), we observe that

$$\left.\begin{array}{r} \lambda_1 + \lambda_2 = p \\ \textit{and } \lambda_1\lambda_2 = -q \end{array}\right\} \tag{5.4}$$

These are explicity given by the formula:

$$\lambda = [p \pm \sqrt{p^2 + 4q}]/2 \tag{5.5}$$

Under our assumptions we have the information that

$$q > 0,\ p \geq 0,\ p + q \leq 1 \tag{5.6}$$

Since $q > 0$, we can use (5.4) to infer that the roots λ_1, λ_2 are real and they are of opposite signs. Without loss of generality, let us denote the positive root by λ_1 and the negative root by λ_2.

Using (5.4), (5.6), we have

$$1 \geq p + q = \lambda_1 + \lambda_2 - \lambda_1\lambda_2 = \lambda_1 + (1 - \lambda_1)\lambda_2$$

so that

$$(1 - \lambda_1) \geq (1 - \lambda_1)\lambda_2 \tag{5.7}$$

Now, if $\lambda_1 > 1$, then we would get $(1-\lambda_1) < 0$, and $(1-\lambda_1)\lambda_2 > 0$ (since $\lambda_2 < 0$) contradicting (5.7). Thus, we can conclude that

$$0 < \lambda_1 \leq 1 \tag{5.8}$$

Now, using (5.4), we have $(-\lambda_2) = \lambda_1 - p \leq \lambda_1 \leq 1$. Thus, neither characteristic root can exceed 1 in absolute value.

5.2 Characteristic Roots Associated with the Ramsey-Euler Equation

Consider the Ramsey-Euler equation:

$$u_3(x_t, x_{t+1}, x_{t+2}) + \delta u_2(x_{t+1}, x_{t+2}, x_{t+3}) + \delta^2 u_1(x_{t+2}, x_{t+3}, x_{t+4}) = 0 \tag{5.9}$$

In particular, of course, $x_{t+s} = x^*$ for $s = 0, 1, 2, 3, 4$ satisfies (5.9):

$$u_3(x^*, x^*, x^*) + \delta u_2(x^*, x^*, x^*) + \delta^2 u_1(x^*, x^*, x^*) = 0 \tag{5.10}$$

If we use the Mean-Value theorem around (x^*, x^*, x^*) to evaluate the difference between the left-hand sides of (5.9) and (5.10), but ignore the second-order terms (so that one obtains a "first-order" or "linear" approximation to the difference) we get (dropping the point of evaulation (x^*, x^*, x^*) to ease the writing) the expression:

$$\delta^2 u_{13}\varepsilon_{t+4} + (\delta^2 u_{12} + \delta u_{23})\varepsilon_{t+3} + (\delta^2 u_{11} + \delta u_{22} + u_{33})\varepsilon_{t+2} + (\delta u_{21} + u_{32})\varepsilon_{t+1} + u_{31}\varepsilon_t$$

If we substitute β^{t+s} for $\varepsilon_{t+s}(s = 0, 1, 2, 3, 4)$, and equate the resulting expression to zero, we get the characteristic equation associated with the Ramsey-Euler equation (5.9):

$$\delta^2 u_{13}\beta^4 + (\delta^2 u_{12} + \delta u_{23})\beta^3 + (\delta^2 u_{11} + \delta u_{22} + u_{33})\beta^2 + (\delta u_{21} + u_{32})\beta + u_{31} = 0 \tag{5.11}$$

The idea is that the roots of this characteristic equation will reflect local behavior around the stationary optimal stock, x^*, of solutions to Ramsey-Euler equations.

We now show that the characteristic roots associated with the optimal policy function, which we analyzed in Section 5.1, must be solutions to the characteristic equation (5.11). By continuity of the optimal policy function, we can choose a neighborhood M of (x^*, x^*) such that for all (x, y) in M, $(y, h(x, y))$, $(h(x, y),\ h(y, h(x, y)))$ and $(h(y, h(x, y)),\ h(h(x, y),\ h(y, h(x, y)))$ are in M', and $(x, y, h(x, y))$, $(y, h(x, y),\ h(y, h(x, y)))$ and $(h(x, y), h(y, h(x, y)),\ h(h(x, y), h(y, h(x, y))))$ are in N. Thus, the Ramsey-Euler equation yields the following *identity* in (x, y):

$$\begin{aligned} W(x, y) &= u_3(x, y, h(x, y)) \\ &\quad + \delta u_2(y, h(x, y),\ h(y, h(x, y))) \\ &\quad + \delta^2 u_1(h(x, y),\ h(y, h(x, y)),\ h(h(x, y),\ h(y, h(x, y)))) \\ &= 0 \end{aligned} \tag{5.12}$$

If we differentiate W with respect to x and evaluate the derivatives of u at (x^*, x^*, x^*), and the derivatives of h at (x^*, x^*), then the derivative $\partial W(x^*, x^*)/\partial x$ must be equal to zero. We can write the derivative (after dropping the points of evaluation (x^*, x^*, x^*) and (x^*, x^*) to ease the writing) as:

$$
\begin{aligned}
\partial W(x^*,x^*)/\partial x &= u_{31} + u_{33}h_1 + \delta[u_{22}h_1 + u_{23}h_2h_1] + \delta^2[u_{11}h_1 + u_{12}h_2h_1 + u_{13}(h_1)^2 + u_{13}(h_2)^2h_1] \\
&= u_{31} + [u_{33} + \delta u_{22} + \delta^2 u_{11}]h_1 + \delta[u_{23} + \delta u_{12}]h_1h_2 + \delta^2 u_{13}[(h_1)^2 + (h_2)^2h_1]
\end{aligned}
$$

Denote $[u_{33} + \delta u_{22} + \delta^2 u_{11}]$ by $\hat{C}$, and $[u_{23} + \delta u_{12}]$ by $\hat{D}$. Then, we have:

$$
\partial W(x^*,x^*)/\partial x = u_{31} + \hat{C}h_1 + \delta\hat{D}h_1h_2 + \delta^2 u_{13}[(h_1)^2 + (h_2)^2h_1] = 0 \tag{5.13}
$$

Similarly, if we differentiate W with respect to y, and evaluate the derivatives of u at (x^*,x^*,x^*), and the derivatives of h at (x^*,x^*), then the derivative $\partial W(x^*,x^*)/\partial y$ must be equal to zero. We can write the derivative (after dropping the points of evaluation (x^*,x^*,x^*) and (x^*,x^*) to ease the writing) as:

$$
\begin{aligned}
\partial W(x^*,x^*)/\partial y &= u_{32} + u_{33}h_2 + \delta[u_{21} + u_{22}h_2 + u_{23}h_1 + u_{23}(h_2)^2] \\
&\quad + \delta^2[u_{11}h_2 + u_{12}h_1 + u_{12}(h_2)^2 + 2u_{13}h_1h_2 + u_{13}(h_2)^3] \\
&= u_{32} + \delta u_{21} + \hat{C}h_2 + \delta\hat{D}h_1 + 2\delta^2 u_{13}h_1h_2 + \delta\hat{D}(h_2)^2 + \delta^2 u_{13}(h_2)^3
\end{aligned}
$$

Rearranging terms yields the derivative:

$$
\partial W(x^*,x^*)/\partial y = (u_{32} + \delta u_{21}) + \hat{C}h_2 + \delta\hat{D}[h_1 + (h_2)^2] + \delta^2 u_{13}[2h_1 + (h_2)^2]h_2 = 0 \tag{5.14}
$$

We recall from Section 5.1 that if λ is a characteristic root associated with the optimal policy function, then $(\lambda)^2 = h_2\lambda + h_1$. Using this information in (5.13) and (5.14), we get:

$$
\begin{aligned}
\partial W(x^*,x^*)/\partial x + \lambda\partial W(x^*,x^*)/\partial y &= u_{31} + [u_{32} + \delta u_{21}]\lambda + \hat{C}(h_1 + h_2\lambda) + \delta\hat{D}[h_1\lambda + (h_2)^2\lambda + h_1h_2] \\
&\quad + \delta^2 u_{13}[(h_1)^2 + (h_2)^2h_1 + 2h_1h_2\lambda + (h_2)^3\lambda] \\
&= u_{31} + [u_{32} + \delta u_{21}]\lambda + \hat{C}\lambda^2 + \delta\hat{D}[h_1\lambda + h_2\lambda^2] \\
&\quad + \delta^2 u_{13}[h_1(h_1 + h_2\lambda) + h_1h_2\lambda + (h_2)^2[h_1 + h_2\lambda]] \\
&= u_{31} + [u_{32} + \delta u_{21}]\lambda + \hat{C}\lambda^2 + \delta\hat{D}\lambda^3 \\
&\quad + \delta^2 u_{13}[h_1\lambda^2 + h_1h_2\lambda + (h_2)^2\lambda^2] \\
&= u_{31} + [u_{32} + \delta u_{21}]\lambda + \hat{C}\lambda^2 + \delta\hat{D}\lambda^3 + \delta^2 u_{13}[h_1\lambda^2 + h_2\lambda^3] \\
&= u_{31} + [u_{32} + \delta u_{21}]\lambda + \hat{C}\lambda^2 + \delta\hat{D}\lambda^3 + \delta^2 u_{13}\lambda^4 = 0
\end{aligned}
$$

This completes the verification of our claim.

We now show how the characteristic roots associated with the optimal policy function (analyzed in Section 5.1) can be found by calculating the characteristic roots of (5.11).

Notice that $\beta = 0$ is *not* a solution to (5.11) since $u_{13} \neq 0$. We can, therefore, use the transformed variable:

$$
\mu = \delta\beta + (1/\beta)
$$

to examine the roots of (5.11). Using this transformation, (5.11) becomes

$$
u_{13}\mu^2 + (\delta u_{12} + u_{23})\mu + [\delta^2 u_{11} + \delta u_{22} + u_{33} - 2\delta u_{13}] = 0 \tag{5.15}
$$

Let us define $\mathbb{G} : \mathbb{R} \to \mathbb{R}$ by

$$\mathbb{G}(\mu) = u_{13}\mu^2 + (\delta u_{12} + u_{23})\mu + [\delta^2 u_{11} + \delta u_{22} + u_{33} - 2\delta u_{13}] \tag{5.16}$$

Since the Hessian of u is negative definite, we have $u_{11} < 0$, $u_{22} < 0$, $u_{33} < 0$, and since $u_{13} > 0$, we have

$$[\delta^2 u_{11} + \delta u_{22} + u_{33} - 2\delta u_{13}]/u_{13} < 0$$

Denoting the roots of (5.15), which is a quadratic in μ, by μ_1 and μ_2, we note that

$$\mu_1\mu_2 < 0 \tag{5.17}$$

so these roots are necessarily real. We denote the positive root by μ_1 and the negative root by μ_2.

Given $\mu_i (i = 1, 2)$, we can obtain the corresponding roots of β by solving the quadratic

$$\delta\beta + (1/\beta) = \mu_i \tag{5.18}$$

We denote the roots of (5.18) corresponding to μ_1 by β_1 and β_2 [with $|\beta_1| = \min[|\beta_1|, |\beta_2|]$] and the roots of (5.18) corresponding to μ_2 by β_3 and β_4 [with $|\beta_3| = \min[|\beta_3|, |\beta_4|]$].

Define the function $\mathbb{F} : \mathbb{R}^2 \to \mathbb{R}$ by

$$\mathbb{F}(\beta; m) = \delta\beta^2 - m\beta + 1 \tag{5.19}$$

Then β_1 and β_2 are the roots of $\mathbb{F}(\beta; \mu_1) = 0$, and β_3 and β_4 are the roots of $\mathbb{F}(\beta; \mu_2) = 0$.

Using our analysis in Section 5.1, we can show that the roots β_1 and β_2 are real, and

$$0 < \beta_1 \leq 1 < \beta_2 \tag{5.20}$$

To see this, recall that λ_1 and λ_2 are solutions of (5.3). These are real and of opposite signs. Thus, examining (5.19), it is clear that λ_1 and λ_2 must correspond to different μ_i. This means that β_1 and β_2 are real, and so are β_3 and β_4.

Now, note that since β_1 and β_2 solve the equation

$$\delta\beta^2 - \mu_1\beta + 1 = 0 \tag{5.21}$$

and $\mu_1 > 0$, we have $\beta_1\beta_2 = (1/\delta) > 0$ and $(\beta_1 + \beta_2) = (\mu_1/\delta) > 0$. Thus, β_1 and β_2 are both positive.

Since β_3 and β_4 are roots of the equation

$$\delta\beta^2 - \mu_2\beta + 1 = 0 \tag{5.22}$$

we have $\beta_3\beta_4 = (1/\delta) > 0$ and $\beta_3 + \beta_4 = (\mu_2/\delta) < 0$. Thus β_3 and β_4 are of the same sign, and they must both be negative.

It follows from the above analysis that λ_1 must be one of the roots β_1 and β_2, and λ_2 must be one of the roots β_3 and β_4. Further, since $\lambda_1 \leq 1$, and $\beta_1\beta_2 = (1/\delta) > 1$, $\lambda_1 = \beta_1$ and $\beta_2 \geq (1/\delta)$. This establishes (5.20).

Similarly, we can show that:

$$0 > \beta_3 \geq -1 > \beta_4 \tag{5.23}$$

Since $(-\lambda_2) \leq 1$, and $\beta_3\beta_4 = (1/\delta) > 1$, $(-\lambda_2) = (-\beta_3)$ and $(-\beta_4) \geq (1/\delta)$. This establishes (5.23).

5.3 Differentiability of the Optimal Policy Function

In Section 5.1, we assumed that the optimal policy function was continuously differentiable in a neighborhood of the steady state, x^*. We used this to obtain the characteristic roots associated with the optimal policy function, and to relate them (in Section 5.2) to the characteristic roots associated with the Ramsey-Euler equation. To complete our analysis, we need to show that the optimal policy function is indeed continuously differentiable in a neighborhood of the steady state, x^*. We do this by applying the Stable Manifold Theorem.[14]

We have seen in Section 5.2 that the characteristic roots $(\beta_1, \beta_2, \beta_3, \beta_4)$ associated with the equation (5.11) satisfy the restrictions:

$$\beta_4 < -1 \leq \beta_3 < 0 < \beta_1 \leq 1 < \beta_2 \tag{5.24}$$

We *assume* now that the generic case in (5.24) holds; that is, the weak inequalities in (5.24) are replaced by strict inequalities:

$$\beta_4 < -1 < \beta_3 < 0 < \beta_1 < 1 < \beta_2 \tag{5.25}$$

We wish to analyze the behavior of the Ramsey-Euler dynamical system near the steady state, x^*. To this end, we define:

$$F(v, w, x, y, z) = u_3(v, w, x) + \delta u_2(w, x, y) + \delta^2 u_1(x, y, z)$$

in a neighborhood $N' \equiv Q^5$ of $(x^*, x^*, x^*, x^*, x^*)$ [where $Q = (x^* - \varepsilon, x^* + \varepsilon)$ and ε is as given in Section 5.1]. Then, F is C^1 on N'. We note that:

$$D_5 F(x^*, x^*, x^*, x^*, x^*) = \delta^2 u_{13}(x^*, x^*, x^*) \neq 0$$

and so we can apply the implicit function theorem[15] to obtain an open set $\tilde{U}$ containing (x^*, x^*, x^*, x^*), and an open set V containing x^*, and a unique function $\Phi : \tilde{U} \to V$, such that:

$$u_3(v, w, x) + \delta u_2(w, x, y) + \delta^2 u_1(x, y, \Phi(v, w, x, y)) = 0 \; for\ all\ (v, w, x, y) \in \tilde{U} \tag{5.26}$$

and:

$$\Phi(x^*, x^*, x^*, x^*) = x^* \tag{5.27}$$

Further, Φ is C^1 on $\tilde{U}$. Clearly, we can pick an open set $\hat{U} \subset \tilde{U}$, with $\hat{U}$ containing (x^*, x^*, x^*, x^*),

[14] The use of stable manifold theory to optimal growth was pioneered by Scheinkman (1976). Since then, it has figured prominently in the theoretical work of Araujo and Scheinkman (1977) and Santos (1991), and in numerous applications of this theory to dynamic macroeconomic models.

[15] See, for example, Rosenlicht (1986), p.205-209.

such that $\Phi(\hat{U}) \subset Q$.

Define the set $U' = \{(v', w', x', y') \in \mathbb{R}^4 : (v', w', x', y') = (v - x^*, w - x^*, x - x^*, y - x^*)$ for some $(v, w, x, y) \in \hat{U}\}$. Thus, U' is a neighborhood of $(0, 0, 0, 0)$, a translation of the set $\hat{U}$ by subtraction of the point (x^*, x^*, x^*, x^*) from each point $(v, w, x, y) \in \hat{U}$. Now, define $G : U' \to \mathbb{R}^4$ as follows:

$$\left.\begin{aligned} G^1(X_1, X_2, X_3, X_4) &= X_2 \\ G^2(X_1, X_2, X_3, X_4) &= X_3 \\ G^3(X_1, X_2, X_3, X_4) &= X_4 \\ G^4(X_1, X_2, X_3, X_4) &= \Phi(x^* + X_1, x^* + X_2, x^* + X_3, x^* + X_4) - x^* \end{aligned}\right\} \tag{5.28}$$

Note that $G(0, 0, 0, 0) = (0, 0, 0, 0)$, using (5.27).

The Ramsey-Euler dynamics near the steady state is governed by (5.26). This gives rise to the (four-dimensional) dynamical system:

$$X_{t+1} = G(X_t) \tag{5.29}$$

In order to apply the standard form of the Stable Manifold Theorem, however, we need to transform the variables appearing in this dynamical system.

To this end, we proceed as follows. Given G, we can calculate the Jacobian matrix of G at $(0, 0, 0, 0)$:

$$J_G(0) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \Phi_1(x^*, x^*, x^*, x^*) & \Phi_2(x^*, x^*, x^*, x^*) & \Phi_3(x^*, x^*, x^*, x^*) & \Phi_4(x^*, x^*, x^*, x^*) \end{bmatrix} \tag{5.30}$$

The entries in the last row of $J_G(0)$ can be related to the second-order derivatives of u at (x^*, x^*, x^*). Differentiating (5.26) with respect to v, w, x, y and evaluating the relevant derivatives at $(v, w, x, y) = (x^*, x^*, x^*, x^*)$, we obtain:

$$\begin{aligned} &u_{31}(x^*, x^*, x^*) + \delta^2 u_{13}(x^*, x^*, x^*)\Phi_1(x^*, x^*, x^*, x^*) = 0 \\ &u_{32}(x^*, x^*, x^*) + \delta u_{21}(x^*, x^*, x^*) + \delta^2 u_{13}(x^*, x^*, x^*)\Phi_2(x^*, x^*, x^*, x^*) = 0 \\ &u_{33}(x^*, x^*, x^*) + \delta u_{22}(x^*, x^*, x^*) + \delta^2 u_{11}(x^*, x^*, x^*) + \delta^2 u_{13}(x^*, x^*, x^*)\Phi_3(x^*, x^*, x^*, x^*) = 0 \\ &\delta u_{23}(x^*, x^*, x^*) + \delta^2 u_{12}(x^*, x^*, x^*) + \delta^2 u_{13}(x^*, x^*, x^*)\Phi_4(x^*, x^*, x^*, x^*) = 0 \end{aligned} \tag{5.31}$$

These equations yield:

$$\left.\begin{aligned} \Phi_1(x^*, x^*, x^*, x^*) &= -(1/\delta^2) \\ \Phi_2(x^*, x^*, x^*, x^*) &= -\frac{[u_{32}(x^*, x^*, x^*) + \delta u_{21}(x^*, x^*, x^*)]}{\delta^2 u_{13}(x^*, x^*, x^*)} \\ \Phi_3(x^*, x^*, x^*, x^*) &= -\frac{[u_{33}(x^*, x^*, x^*) + \delta u_{22}(x^*, x^*, x^*) + \delta^2 u_{11}(x^*, x^*, x^*)]}{\delta^2 u_{13}(x^*, x^*, x^*)} \\ \Phi_4(x^*, x^*, x^*, x^*) &= -\frac{[\delta u_{23}(x^*, x^*, x^*) + \delta^2 u_{12}(x^*, x^*, x^*)]}{\delta^2 u_{13}(x^*, x^*, x^*)} \end{aligned}\right\} \tag{5.32}$$

Define the Vandermonde matrix:

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \beta_1 & \beta_3 & \beta_2 & \beta_4 \\ \beta_1^2 & \beta_3^2 & \beta_2^2 & \beta_4^2 \\ \beta_1^3 & \beta_3^3 & \beta_2^3 & \beta_4^3 \end{bmatrix} \tag{5.33}$$

Note that the unusual order in the Vandermonde matrix is to be explained by the fact that the characteristic roots β_1 and β_3 are less than one in absolute value, while β_2 and β_4 are greater than one in absolute value. [This order becomes important in the application of the Stable Manifold Theorem below]. Define the diagonal matrix of characteristic values:

$$\mathbb{B} = \begin{bmatrix} \beta_1 & 0 & 0 & 0 \\ 0 & \beta_3 & 0 & 0 \\ 0 & 0 & \beta_2 & 0 \\ 0 & 0 & 0 & \beta_4 \end{bmatrix} \tag{5.34}$$

Now, denoting by $\mathbb{A}$ the Jacobian matrix $J_G(0)$, we can verify (using (5.30),(5.32) and (5.11)) that:

$$\mathbb{A}P = P\mathbb{B} = \begin{bmatrix} \beta_1 & \beta_3 & \beta_2 & \beta_4 \\ \beta_1^2 & \beta_3^2 & \beta_2^2 & \beta_4^2 \\ \beta_1^3 & \beta_3^3 & \beta_2^3 & \beta_4^3 \\ \beta_1^4 & \beta_3^4 & \beta_2^4 & \beta_4^4 \end{bmatrix} \tag{5.35}$$

This means that $(\beta_1, \beta_3, \beta_2, \beta_4)$ are the characteristic roots of $\mathbb{A}$, with the column vectors of P constituting a set of characteristic vectors of $\mathbb{A}$, corresponding to these characteristic roots. The Vandermonde matrix is known to be non-singular[16], so we get the spectral decomposition:

$$P^{-1}\mathbb{A}P = \mathbb{B} \tag{5.36}$$

Returning now to our dynamical system (5.29), we rewrite it as:

$$X_{t+1} = \mathbb{A}X_t + [G(X_t) - \mathbb{A}X_t] \tag{5.37}$$

Multiplying through in (5.37) by P^{-1}, we obtain:

$$P^{-1}X_{t+1} = (P^{-1}\mathbb{A}P)P^{-1}X_t + [P^{-1}G(PP^{-1}X_t) - (P^{-1}\mathbb{A}P)P^{-1}X_t] \tag{5.38}$$

Thus, using (5.36), and defining new variables $Y = P^{-1}X$, we get:

$$Y_{t+1} = \mathbb{B}Y_t + [P^{-1}G(PY_t) - \mathbb{B}Y_t] \tag{5.39}$$

Denote by U the set $\{Y : Y = P^{-1}X \text{ for some } X \in U'\}$, and define $g : U \to \mathbb{R}^4$ as follows:

$$g(Y) = P^{-1}G(PY) - \mathbb{B}Y \tag{5.40}$$

[16] Several methods are known for computing the inverse of a Vandermonde matrix. For one such approach, see Parker (1964).

Note that by (5.28), we have:

$$g(0,0,0,0) = (0,0,0,0) \tag{5.41}$$

Also, we obtain by differentiating (5.40) and evaluating the derivatives at $(0,0,0,0)$:

$$J_g(0) = P^{-1}J_G(P0)P - \mathbb{B} = P^{-1}J_G(0)P - \mathbb{B} = P^{-1}\mathbb{A}P - \mathbb{B} = 0 \tag{5.42}$$

Thus, the dynamical system (5.39) can now be written as :

$$Y_{t+1} = \mathbb{B}Y_t + g(Y_t) \tag{5.43}$$

with $g(0) = 0$ and $J_g(0) = 0$.

The Stable Manifold Theorem can be applied to the dynamical system (5.43). We give below the particular statement of it (due to Irwin (1970)) that is directly applicable[17].

Stable Manifold Theorem for a Fixed Point (Irwin):
Let $E = E_1 \times E_2$ be a Banach Space and let $T_1 : E_1 \to E_1$ and $T_2 : E_2 \to E_2$ be isomorphisms with $\max\{\|T_1\|, \|T_2^{-1}\|\} < 1$. Let U be an open neighborhood of 0 in E and let $g : U \to E$ be a C^r map $(r \geq 1)$ with $g(0) = 0$ and $Dg(0) = 0$. Let $f = T_1 \times T_2 + g$. Then, there exist open balls C and D centered at 0 in E_1 and E_2 respectively, and a unique map $H : C \to D$ such that $f(graph(H)) \subset graph(H)$. The map H is C^r on the open ball C and $DH(0) = 0$. Further, for all $z \in C \times D$, $f^n(z) \to 0$ as $n \to \infty$ if and only if $z \in graph(H)$).

To apply the theorem, we define the maps $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ as follows:

$$T_1(z) = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} ; T_2(z) = \begin{bmatrix} \beta_2 & 0 \\ 0 & \beta_4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Note that:

$$T_2^{-1}(z') = \begin{bmatrix} (1/\beta_2) & 0 \\ 0 & (1/\beta_4) \end{bmatrix} \begin{bmatrix} z_1' \\ z_2' \end{bmatrix}$$

so that, using (5.25), we have $\|T_1\| < 1$ and $\|T_2^{-1}\| < 1$. Applying the theorem in our context (with $r = 1$) we get the C^1 function H with the properties stated above. We wish to conclude from this that the policy function, h, is C^1 in a neighborhood of (x^*, x^*).

First, we note that $H(0,0) = (0,0)$. To see this, we check that $f(0,0,0,0) = g(0,0,0,0) = (0,0,0,0)$ by (5.41), so that $f^n(0,0,0,0) = (0,0,0,0)$, and so by the Stable Manifold Theorem, $(0,0,0,0) \in graph(H)$. That is, $H(0,0) = (0,0)$.

Next, we define a function, $K : \mathbb{R}^2 \times \mathbb{R}^2 \times C \to \mathbb{R}^4$ as follows:

$$K(a,b,z) = P^{-1}(a,b) - (z, H(z)) \tag{5.44}$$

Clearly, K is C^1 on its domain, and $K(0,0,0,0,0,0) = (0,0,0,0)$, since $H(0,0) = (0,0)$. Further, the matrix $(D_jK^i(0,0,0,0,0,0))$, where $i = 1,2,3,4$ and $j = 3,4,5,6$ can be checked to be

[17] A good exposition of Irwin's result can be found in Franks (1979).

non-singular. To see this, denote P^{-1} by R, and write R as follows:

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

where each R_{ij} (with $i = 1, 2; j = 1, 2$) a 2×2 matrix. Then, we have:

$$(D_j K^i(0,0,0,0,0,0)) = \begin{bmatrix} R_{12} & -I \\ R_{22} & 0 \end{bmatrix}$$

where I is the 2×2 identity matrix, and 0 is the 2×2 null matrix. Thus, the matrix $(D_j K^i(0,0,0,0,0,0))$ is non-singular if and only if R_{22} is non-singular. To verify that R_{22} is non-singular, we write (by definition of R) :

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

where each P_{ij} (with $i = 1, 2; j = 1, 2$) is a 2×2 sub-matrix of P. This yields the equations:

$$\left.\begin{array}{l} R_{21}P_{11} + R_{22}P_{21} = 0 \\ R_{21}P_{12} + R_{22}P_{22} = I \end{array}\right\} \tag{5.45}$$

Clearly, P_{11} is non-singular, since $\det(P_{11}) = \beta_3 - \beta_1 < 0$ (by (5.25)). Thus, $R_{21} = -R_{22}P_{21}P_{11}^{-1}$ (from the first equation of (5.45)) and using this in the second equation of (5.45), we obtain $R_{22}[P_{22} - P_{21}P_{11}^{-1}P_{12}] = I$. This establishes that R_{22} is non-singular.

We can now use the implicit function theorem to obtain an open set $E' \subset \mathbb{R}^2$ containing $(0,0)$, an open set $C' \subset C$ containing $(0,0)$, and an open set $E'' \subset \mathbb{R}^2$ containing $(0,0)$, and unique functions $L^1 : E' \to E''$ and $L^2 : E' \to C'$, such that:

$$K(a, L^1(a), L^2(a)) = 0 \; for \; all \; a \in E' \tag{5.46}$$

and:

$$L^1(0,0) = (0,0); \;\; L^2(0,0) = (0,0) \tag{5.47}$$

Further, L^1 and L^2 are C^1 on E'. Using the definition of K, we have from (5.46) :

$$P^{-1}(a, L^1(a)) = (L^2(a), H(L^2(a))) \; for \; all \; a \in E' \tag{5.48}$$

Now, we look at the optimal policy function, h. Pick $0 < \varepsilon' < \varepsilon$ (where ε is given as in Section 5.1) so that $(-\varepsilon', \varepsilon')^4 \subset U'$ (where U' is given as in (5.28)), and $P^{-1}z \in C \times D$ for all $z \in (-\varepsilon', \varepsilon')^4$. Denote $(-\varepsilon', \varepsilon')$ by S.

Pick any $(z_1, z_2) \in S^2$. Define $(x_1, x_2) = (x^*, x^*) + (z_1, z_2)$. Then the sequence $\{x_t\}$ satisfying $x_{t+2} = h(x_t, x_{t+1})$ for $t \geq 1$ is well-defined and $x_t \to x^*$ as $t \to \infty$. Thus, the sequence $\{z_t\}$ satisfying $z_t = x_t - x^*$ for $t \geq 1$ is well-defined and $z_t \to 0$ as $t \to \infty$. Further, since $(z_1, z_2) \in S^2$, we have $z_t \in S$ for all $t \geq 1$ (by the proof of Theorem 2). Then, we have:

$$(z_t, z_{t+1}, z_{t+2}, z_{t+3}) \in U' \; for \; t \geq 1 \tag{5.49}$$

and:

$$P^{-1}(z_t, z_{t+1}, z_{t+2}, z_{t+3}) \in C \times D \ for \ t \geq 1 \tag{5.50}$$

Using (5.28) and (5.49), we can write for $t \geq 1$,

$$f(P^{-1}(z_t, z_{t+1}, z_{t+2}, z_{t+3})) = P^{-1}(z_{t+1}, z_{t+2}, z_{t+3}, \Phi(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3}) - x^*) \tag{5.51}$$

Using (5.12), we have for $t \geq 1$,

$$\begin{aligned} u_3(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}) + \delta u_2(x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3}) \\ + \delta^2 u_1(x^* + z_{t+2}, x^* + z_{t+3}, x^* + z_{t+4}) = 0 \end{aligned} \tag{5.52}$$

Using (5.26) and (5.49), we have for $t \geq 1$,

$$\begin{aligned} u_3(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}) + \delta u_2(x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3}) \\ + \delta^2 u_1(x^* + z_{t+2}, x^* + z_{t+3}, \Phi(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3})) = 0 \end{aligned} \tag{5.53}$$

Note that by (5.49), $\Phi(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3}) \in Q$. Since $u_{13} > 0$ on Q^3, (5.52) and (5.53) yield (by an application of the Mean Value theorem):

$$\Phi(x^* + z_t, x^* + z_{t+1}, x^* + z_{t+2}, x^* + z_{t+3}) = x^* + z_{t+4} \tag{5.54}$$

Using (5.54) in (5.51), we obtain:

$$f(P^{-1}(z_t, z_{t+1}, z_{t+2}, z_{t+3})) = P^{-1}(z_{t+1}, z_{t+2}, z_{t+3}, z_{t+4}) \ for \ t \geq 1 \tag{5.55}$$

We can infer from (5.55) that:

$$f^n(P^{-1}(z_1, z_2, z_3, z_4)) = P^{-1}(z_{n+1}, z_{n+2}, z_{n+3}, z_{n+4}) \ for \ n \geq 1 \tag{5.56}$$

Since the right hand-side of (5.56) converges to $(0,0,0,0)$ as $n \to \infty$, we must have $f^n(P^{-1}(z_1, z_2, z_3, z_4)) \to (0,0,0,0)$ as $n \to \infty$. By the Stable Manifold Theorem, then, we must have:

$$P^{-1}(z_1, z_2, z_3, z_4) \in graph(H) \tag{5.57}$$

Define a function $\psi : S^2 \to \mathbb{R}^2$ by:

$$\psi(z_1, z_2) = (h(x^* + z_1, x^* + z_2) - x^*, h(x^* + z_2, h(x^* + z_1, x^* + z_2)) - x^*) \ for \ all \ z \in S^2$$

Then $\psi(0,0) = (0,0)$ and (5.57) shows that, given any $z = (z_1, z_2) \in S^2$, we must have $P^{-1}(z, \psi(z)) \in graph(H)$. Thus, given any $z \in S^2$, there is $z' \in C$, such that:

$$P^{-1}(z, \psi(z)) = (z', H(z'))$$

Clearly, such a z' must be unique. Thus, there is a function, $\mathbb{K} : S^2 \to C$ such that:

$$P^{-1}(z, \psi(z)) = (\mathbb{K}(z), H(\mathbb{K}(z))) \ for \ all \ z \in S^2 \tag{5.58}$$

Note that since $\psi(0,0) = (0,0)$, (5.58) implies that $\mathbb{K}(0,0) = (0,0)$. Defining $S' = S^2 \cap E'$, we have from (5.58),

$$P^{-1}(z, \psi(z)) = (\mathbb{K}(z), H(\mathbb{K}(z))) \; for \; all \; z \in S' \tag{5.59}$$

On the other hand, from (5.48), we have:

$$P^{-1}(z, L^1(z)) = (L^2(z), H(L^2(z))) \; for \; all \; z \in S' \tag{5.60}$$

Since L^1 and L^2 are the unique functions satisfying (5.60) and (5.47), and since $\psi(0,0) = (0,0)$ and $\mathbb{K}(0,0) = (0,0)$, we must have $\psi = L^1$ and $\mathbb{K} = L^2$ on S'. Since L^1 is C^1 on S', we can conclude that ψ is C^1 on S'. Using the definition of ψ , it follows that the optimal policy function, h, is C^1 on S'.

6 Concluding Remarks

The purpose of the paper was to complete the program sketched in the contribution of Samuelson (1971), by providing both a complete local and global analysis of the model under which the standard results of the Ramsey model continue to hold even with dependence of tastes between periods. We approached the problem by trying to identify structures in the *reduced form* of the model under which this would be true. The reduced form model involves a utility function which depends on the values of the state variable (capital stock) at *three* successive dates (instead of the usual two).We showed that supermodularity of the reduced form utility function (in the three variables), and a single-crossing condition provides such a structure. The methods used indicate that our results should generalize to situations in which the reduced form utility function depends on the values of the state variable in more than three periods (which correspond to situations in which utility function used to evaluate current consumption depends on several periods of past consumption).

We examined the implication of this structure for the model of Samuelson (1971) with intertemporal dependence of tastes. Our analysis indicated the conditions on the primitive form of the model under which the assumptions on the corresponding reduced form are met. It also indicated plausible scenarios in which the stated assumptions on the reduced-form *would not* be satisfied. Thus, identifying these assumptions provides a good handle on the richer dynamics that this model can generate when these assumptions fail; exploration of this topic is undertaken in Mitra and Nishimura (2001).

Application of our methods to models of habit formation is a natural direction of enquiry. The model of Boyer (1978) on habit formation, where utility is assumed to be increasing both in current and in past consumption, can be accommodated in our framework. Other frameworks of habit formation, where utility is increasing in present consumption, but decreasing in past consumption, are ruled out by the basic assumptions of our model. However, we indicated with an example that our methods and results are valid in *some* of these frameworks of habit formation as well. Models of habit formation leading to non-concave utility functions cannot be directly addressed by the methods of this paper, and constitute a potentially interesting area of future research.

References

[1] A.B. Abel (1990), Asset Prices under Habit Formation and Catching up with the Joneses, *American Economic Review Papers and Proceedings*, 38-42.

[2] R. Amir (1996), "Sensitivity Analysis of Multisector Optimal Economic Dynamics", *Journal of Mathematical Economics*, 25, 123-141.

[3] R. Amir, L.J. Mirman and W.R. Perkins (1991), "One-Sector Nonclassical Optimal Growth: Optimality Conditions and Comparative Dynamics", *International Economic Review*, 32, 625-792.

[4] A. Araujo (1991), "The Once but not Twice Differentiability of the Policy Functions", *Econometrica*, 59, 1381-1393.

[5] A. Araujo and J.A. Scheinkman (1977), "Smoothness, Comparative Dynamics, and the Turnpike Property", *Econometrica*, 45, 601-620.

[6] R. Becker and J. Boyd (1997), *Capital Theory, Equilibrium Analysis, and Recursive Utility*, Basil Blackwell: Oxford.

[7] J. Benhabib and K. Nishimura (1985), "Competitive Equilibrium Cycles", *Journal of Economic Theory*, 35, 284-306.

[8] B.D. Bernheim and D. Ray (1987), "Economic Growth with Intergenerational Altruism", *Review of Economic Studies*, 54, 227-243.

[9] M.Boyer (1978), "A Habit Forming Optimal Growth Model", *International Economic Review*, 19, 585-609.

[10] C.W. Clark (1976), "A Delayed-Recruitment Model of Population Dynamics, with an Application to Baleen Whale Populations", *Journal of Mathematical Biology*, 3, 381-391.

[11] S. Chakravarty and A.S. Manne (1968), "Optimal Growth when the Instantaneous Utility Function Depends Upon the Rate of Change of Consumption", *American Economic Review*, 58, 1351-1354.

[12] P. Dasgupta (1974), "On Some Problems Arising from Professor Rawls' Conception of Distributive Justice", *Theory and Decision*, 4, 325-344.

[13] A. Deaton (1992), *Consumption*, Oxford University Press.

[14] J. Franks (1979), "Manifolds of C^r Mappings and Applications to Differentiable Dynamical Systems", *Studies in Analysis*, Gian-Carlo Rota (editor), New York: Academic Press, 271-290.

[15] D. Gale (1967), "On Optimal Development in a Multi-Sector Economy", *Review of Economic Studies*, 34, 1-18.

[16] M.L.J. Hautus and Bolis, T.S.(1979), "A Recursive Real Sequence", *American Mathematical Monthly*, 86, 865-866.

[17] G.M. Heal and H.E. Ryder (1973), "Optimal Growth with Intertemporally Dependent Preferences", *Review of Economic Studies*, 40, 1-31.

[18] M.C. Irwin (1970), "On the Stable Manifold Theorem", *Bulletin of the London Mathematical Society*, 2, 196-198.

[19] K. Iwai (1972), "Optimal Economic Growth and Stationary Ordinal Utility: A Fisherian Approach", *Journal of Economic Theory*, 5, 121-151.

[20] E. Kohlberg (1976), "A Model of Economic Growth with Altrusim Between Generations", *Journal of Economic Theory*, 13, 1-13.

[21] T.C. Koopmans (1960), "Stationary Ordinal Utility and Impatience", *Econometrica,* 28, 287-309.

[22] T.C. Koopmans, P.A. Diamond, and R.E. Williamson (1964), "Stationary Ordinal Utility and Time Perspective", *Econometrica,* 32, 82-100.

[23] J. Lane and T. Mitra (1981), "On Nash Equilibrium Programs of Capital Accumulation Under Altruistic Preferences", *International Economic Review*, 22, 309-331.

[24] T.Mitra and K.Nishimura (2001), "Cyclical and Chaotic Optimal Paths in a Model with Intertemporal Complementarity", mimeo.

[25] T. Mitra and K. Nishimura (2003), "Simple Dynamics in a Model of Habit Formation", mimeo.

[26] L. Montrucchio (1998), "Thompson Metric, Contraction Property and Differentiability of Policy Functions", *Journal of Economic Behavior and Organization*, 33, 449-466.

[27] F.D. Parker (1964), "Inverses of Vandermonde Matrices", *American Mathematical Monthly*, 71, 410-411.

[28] M. Rosenlicht (1986), *Introduction to Analysis*, Dover: New York.

[29] S.M. Ross (1983), *Introduction to Stochastic Dynamic Programming*, Academic Press: New York.

[30] P.A. Samuelson (1971), "Turnpike Theorems even though Tastes are Intertemporally Dependent", *Western Economic Journal*, 9, 21-26.

[31] M.S. Santos (1991), "Smoothness of the Policy Function in Discrete-Time Economic Models", *Econometrica*, 59, 1365-1382.

[32] J.A. Scheinkman (1976), "On Optimal Steady States of n-sector Growth Models when Utility is Discounted", *Journal of Economic Theory*, 12, 11-20.

[33] S.M. Sundaresan (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth", *Review of Financial Studies*, 2, 73-89.

[34] D. Topkis (1968), "Minimizing a Submodular Function on a Lattice", *Operations Research*, 26, 305-321.

[35] D. Topkis (1998), *Supermodularity and Complementarity*, Princeton University Press.

[36] H.Y. Wan (1970), "Optimal Savings Programs Under Intertemporally Dependent Preferences", *International Economic Review*, 11, 521-547.