The Japanese Economic Review
Vol. 50, No. 4, December 1999

ON THE EXISTENCE OF CHAOTIC POLICY FUNCTIONS IN DYNAMIC OPTIMIZATION*

By TAPAN MITRA† and GERHARD SORGER‡
†Cornell University ‡University of Vienna

Applying the characterization results from Mitra and Sorger (1999), we show that topological chaos is a robust phenomenon in standard aggregative growth models even under arbitrary mild discounting. Furthermore, we state exact discount factor restrictions, under which two of the most popular examples of chaotic dynamics, the logistic map and the tent map, can be optimal policy functions of aggregative growth models.
JEL Classification Numbers: C61, O41.

1. Introduction

We are concerned with solutions of dynamic optimization problems of the form

$$V(x) = \sup \sum_{t=0}^{+\infty} \rho^t U(x_t, x_{t+1})$$

$$\text{s.t. } (x_t, x_{t+1}) \in \Omega, \qquad t \in \{0, 1, 2, \ldots\}, x_0 = x,$$

where $\rho \in (0, 1)$ is the discount factor, $X \subset \mathbb{R}$ is a compact interval, $\Omega \subseteq X \times X$ is a closed and convex transition possibility set, $U\colon \Omega \mapsto \mathbb{R}$ is a continuous and concave utility function and $x \in X$ is the initial state of the system. Models of this form arise in many different areas of economics, notably in optimal growth theory (see Stokey and Lucas, 1989, and McKenzie, 1986). It is well known that the optimal paths of this problem (starting at any initial state x) are characterized by a continuous function h which maps the state at time t, x_t, to its unique optimal successor state $x_{t+1} = h(x_t)$. Mitra and Sorger (1999) have recently characterized the set of pairs (h, V) which can be the optimal policy function and optimal value function of a problem (Ω, U, ρ) satisfying the standard convexity and continuity assumptions. In the present note, we apply these characterization results in order to:

- investigate the possibility that optimal policy functions are infinitely steep;
- demonstrate that topological chaos is a robust phenomenon in the class of optimization models under consideration even under arbitrary mild discounting; and
- derive exact discount factor restrictions under which two of the most popular examples of chaotic dynamics, the logistic map and the tent map, can be the optimal policy functions of such a model.

All of these issues have already been dealt with in the literature, as will be briefly summarized below. The contribution of the present note is that we use a novel approach

* Gerhard Sorger acknowledges support from the Austrian Science Foundation under grants P10850-SOZ and F010.

Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK.

for analysing these questions which allows us to derive either more complete answers, or the already known answers in a different way.

Many policy functions arising from the one-sector neoclassical model (in which $X = [0, \bar{x}]$) violate Lipschitz continuity at $x = 0$. Mitra and Sorger (1999) present an example in which the optimal policy function has slope $-\infty$ at an interior fixed point. On the other hand, it has been noted by a number of authors that an optimal policy function cannot have the slope $+\infty$ at an interior fixed point (see e.g. Hewage and Neumann, 1990; Mitra, 1996; or Sorger, 1995). We derive this result as an application of the characterization results from Mitra and Sorger (1999). Moreover, we prove by an example that one cannot rule out that an optimal policy function has slope $+\infty$ at an interior point that is not a fixed point.

The possibility of chaotic dynamics in models of the form described above has been known since Boldrin and Montrucchio (1986) and Deneckere and Pelikan (1986). In the examples that they used to demonstrate this result, the discount factors were chosen extremely small (close to $1/100$), which led some researchers to believe that chaos is not after all possible for "reasonable" parameter values. This conjecture was proved wrong by Nishimura *et al.* (1994) and Nishimura and Yano (1995), who demonstrated that chaotic optimal solutions can be found in this class of models even when the discount factor is chosen arbitrarily close to 1. Nishimura *et al.* (1998) provided an alternative proof of the same result which does not rely on boundary solutions. All of these studies have used parametric examples to make their point, and robustness of optimal chaos has been demonstrated by showing that the construction has the desired properties for an open set of parameter values.

In the present paper we derive the existence of chaotic dynamics in optimal growth models as a straightforward implication of the fact (proved in Mitra and Sorger, 1999) that every Lipschitz continuous function with Lipschitz constant L can be an optimal policy function, provided the discount factor is smaller than $1/L^2$. We then show in a general setting that the optimal policy function depends continuously on the utility function and the discount factor. Together with the lower semi-continuity of topological entropy, this establishes the robustness of optimal chaos without the need for a parametrization.

One example that has been considered by many contributors to the literature on the optimality of chaos is the logistic map $h(x) = 4x(1-x)$ defined on the unit interval. It is one of the standard examples of chaotic dynamics. Both Boldrin and Montrucchio (1986) and Deneckere and Pelikan (1986), for example, used this map to prove their results. (As mentioned above, the discount factors had to be chosen close to $1/100$.) Various upper bounds for the set of discount factors that are consistent with the optimality of the logistic map were derived in Mitra (1996), Montrucchio (1994) and Sorger (1992a,b). In particular, Montrucchio (1994) proved that the logistic map cannot be optimal for discount factors larger than $1/16$ if the optimization problem satisfies a certain regularity assumption. In Section 5 we strengthen this result and show that the logistic map can be the optimal policy function of a regular dynamic optimization problem if and only if the discount factor does not exceed $1/16$.

Another standard example of chaotic dynamics is the tent map $h(x) = 1 - |2x - 1|$ (also defined on the unit interval). In contrast to the logistic map, the tent map is not differentiable, so that the approach of Boldrin and Montrucchio (1986) cannot be used to rationalize it. In Sorger (1992b), however, it was shown that the tent map can be rationalized for discount factors smaller than $1/4$; and in Sorger (1994) it was proved that it cannot be rationalized by any dynamic optimization problem with a discount factor greater than or equal to $1/\sqrt{6}$. In the present paper we improve these results by showing that the tent map can be rationalized if and only if $\rho \leqslant 1/4$.

The paper is organized as follows. Section 2 formulates the dynamic optimization problem under consideration and states the assumptions. Section 3 states and discusses the characterization results from Mitra and Sorger (1999) that are the basis of the present note; it contains a result as well as an example concerning the possibility of infinitely steep optimal policy functions. Section 4 deals with the robustness of topological chaos under mild discounting, and Section 5 presents exact discount factor restrictions for the optimality of the logistic map and the tent map.

2. Dynamic optimization problems

Time is measured in discrete periods $t \in \{0, 1, 2, \ldots\}$. At each time t the state of the economic system is described by a vector $x_t \in X$ where the state space $X \subseteq \mathbb{R}$ is a compact interval with non-empty interior. The problem under consideration is to maximize

$$\sum_{t=0}^{+\infty} \rho^t U(x_t, x_{t+1}) \tag{1}$$

over the set of all sequences $(x_t)_{t=0}^{+\infty}$ satisfying the constraints

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

$$x_0 = x. \tag{3}$$

The notation has the following interpretation: ρ is the discount factor, U is the utility function, Ω is the constraint set and $x \in X$ is the initial state. The following assumptions will be used in this paper.

A1. $\Omega \subseteq X \times X$ is a closed and convex set such that the x-section $\Omega_x = \{y \in X | (x, y) \in \Omega\}$ is non-empty for all $x \in X$ and the set $\bigcup_{x \in X} \Omega_x$ has non-empty interior.
A2. $U : \Omega \mapsto \mathbb{R}$ is a continuous and concave function.
A3. $\rho \in (0, 1)$.

We shall refer to the dynamic optimization problem (1)–(3) as problem (Ω, U, ρ). Note that we do not include the initial state x in the description of the problem. Thus, problem (Ω, U, ρ) requires finding the optimal state trajectories from *any* initial state $x \in X$. Assumptions A1–A3 are standard assumptions in the relevant literature, and they imply that the Bellman equation,

$$V(x) = \max\{U(x, y) + \rho V(y) | y \in \Omega_x\},$$

holds for all $x \in X$. Moreover, a path $(x_t)_{t=0}^{+\infty}$ satisfying (2) and (3) is optimal if and only if $V(x_t) = U(x_t, x_{t+1}) + \rho V(x_{t+1})$ holds for all $t \in \{0, 1, 2, \ldots\}$. In general, optimal paths for (1)–(3) need not be unique. To ensure uniqueness, one has to add strict concavity.

A4. The optimal value function V is strictly concave.

If A1–A4 hold, then the following is true. For every $x \in X$, there exists exactly one $y \in \Omega_x$ such that $V(x) = U(x, y) + \rho V(y)$. In other words, there exists a unique maximizer on the right-hand side of the Bellman equation. Let $h(x)$ denote this maximizer; that is,

$$h(x) = \text{argmax}\{U(x, y) + \rho V(y) | y \in \Omega_x\}.$$

The function $h : X \mapsto X$ defined in that way is called the optimal policy function of the optimization problem (Ω, U, ρ). It maps any state $x \in X$ to its optimal successor state $h(x)$. Optimal paths are uniquely determined as the trajectories of the difference equation $x_{t+1} = h(x_t)$ with (3) as the initial condition.

Some of our results involve a concavity assumption that is based on the notion of α-concavity and α-convexity. If α is any real number, then V is α-concave if $x \mapsto V(x) + (\alpha/2)\|x\|^2$ is a concave function. Analogously, we say that V is α-convex if $x \mapsto V(x) - (\alpha/2)\|x\|^2$ is convex.

A5. There exist positive real numbers α and β such that V is α-concave and $(-\beta)$-convex.

It is obvious that this assumption is stronger than A4. We call an optimization problem (Ω, U, ρ) *regular* if it satisfies A1–A5.

In many cases (especially in optimal growth theory) the optimization problem (Ω, U, ρ) is also assumed to satisfy the following monotonicity assumption.

A6. If $x \leqslant \bar{x}$ then $\Omega_x \subseteq \Omega_{\bar{x}}$. The function $x \mapsto U(x, y)$ is non-decreasing and the function $y \mapsto U(x, y)$ is non-increasing.

It is known that an optimization problem (Ω, U, ρ) that satisfies A1–A4 and A6 has a non-decreasing optimal value function V (see e.g. Stokey and Lucas, 1989, Theorem 4.7).

3. The characterization results

The main result from Mitra and Sorger (1999), specialized to a one-dimensional state space, is stated in the following proposition.

Proposition 1. *Let $h : X \mapsto X$ and $V : X \mapsto \mathbb{R}$ be two given functions.*
(i) If there exists a dynamic optimization problem (Ω, U, ρ) on X such that Assumptions A1–A4 hold and such that h is the optimal policy function and V the optimal value function, then the following is true: h and V are continuous; V is strictly concave; and, for all $x \in X$ such that $\partial V(x) \neq \emptyset$ and all $p_x \in \partial V(x)$, there exists $q_x \in \partial V(h(x))$ such that

$$V(h(x)) - V(h(y)) + q_x[h(y) - h(x)] \leqslant (1/\rho)[V(x) - V(y) + p_x(y - x)] \qquad (4)$$

holds for all $y \in X$.
(ii) There exists a dynamic optimization problem (Ω, U, ρ) on X such that A1–A4 hold and such that h is the optimal policy function and V the optimal value function, provided that the following is true: h and V are continuous; V is strictly concave; and, for every $x \in X$, there exist subgradients $p_x \in \partial V(x)$ and $q_x \in \partial V(h(x))$ such that (4)

holds for all $y \in X$. *If, in addition,* V *is non-decreasing, then* (Ω, U, ρ) *can be chosen such that A6 holds.*

Part (i) of this proposition states a condition that is necessary for the pair (h, V) to be the optimal solution of a dynamic optimization problem, whereas part (ii) provides a sufficient condition. The only important difference between these two conditions is that the sufficient condition requires that V is sub-differentiable also at the two boundary points of X.[1]

Remark 1. The statement of the necessary condition in Mitra and Sorger (1999, Theorem 1) uses a slightly different order of the quantifiers. It is clear from the proof in Mitra and Sorger (1999), however, that the version used in Proposition 1(i) above holds. The statement of the sufficient condition in Mitra and Sorger (1999, Theorem 2) requires that V can be extended as a concave function to some open set containing X. In the one-dimensional framework of the present paper, it is easy to check that the sub-differentiability of V at all $x \in X$ is equivalent to that requirement.

If V is sub-differentiable at both x and y, then we can reverse the roles of x and y in (4). Adding (4) to the inequality that is obtained by interchanging x and y in (4), we obtain

$$(q_x - q_y)[h(y) - h(x)] \leqslant (1/\rho)(p_x - p_y)(y - x). \tag{5}$$

This is a more symmetric (but weaker) necessary condition for the rationalizability of the pair (h, V).

Mitra and Sorger (1999) derive the following important corollary from the above result.

Proposition 2. *Let* $h : X \mapsto X$ *be a Lipschitz-continuous function with Lipschitz constant L. For every* $\rho \leqslant 1/L^2$, *there exists an optimization problem* (Ω, U, ρ) *satisfying A1–A6 which has h as its optimal policy function.*

Proposition 1(i) together with Proposition 2 implies that the set of all functions that can be optimal policy functions of a problem satisfying A1–A4 is contained in the set of continuous functions but contains the set of Lipschitz-continuous functions. Therefore, we have a rather good characterization of this set. Getting a better (or even complete) characterization of all possible optimal policy functions seems to be a subtle problem. On the one hand, there are continuous functions which cannot be optimal policy functions, and on the other, there are optimal policy functions that are not Lipschitz-continuous. In the remainder of this section we demonstrate these two properties by applying Proposition 1.

Theorem 1. *Assume that* $h : X \mapsto X$ *is a continuous function which has the fixed point* $x = h(x) \in X$. *Assume that there exists* $z \in \operatorname{int} X$ *such that* $x = h(z)$. *If*

$$\limsup_{y \to x} \frac{h(y) - h(x)}{y - x} = +\infty, \tag{6}$$

then h cannot be the optimal policy function of an optimization problem (Ω, U, ρ) *satisfying A1–A4.*

1) Every concave function is automatically sub-differentiable at any point in the interior of its domain.

Proof. Assumption (6) implies that there exists a sequence $(y_i)_{i=1}^{+\infty}$ such that $\lim_{i\to+\infty} y_i = x$ and $\lim_{i\to+\infty}[h(y_i) - x]/(y_i - x) = +\infty$. From this sequence one can extract a subsequence (again denoted by $(y_i)_{i=1}^{+\infty}$) such that either $h(y_i) \geqslant x$ holds for all i or $h(y_i) \leqslant x$ holds for all i. We consider only the first case; the second one can be dealt with by an analogous argument.

Assume that h is optimal for (Ω, U, ρ) and let V be the corresponding optimal value function. Since $z \in \operatorname{int} X$, we know that V is sub-differentiable at z. Proposition 1(i) implies that V is also subdifferentiable at $x = h(z)$. Let $p_x = \inf\{p | p \in \partial V(x)\}$, and let q_x be an arbitrary element of $\partial V(x)$. Therefore, p_x and q_x are subgradients of V at $x = h(x)$ and

$$(q_x - p_x)[h(y_i) - h(x)] \geqslant 0 \tag{7}$$

for all i.[2]

Define $\lambda_i = (y_i - x)/[h(y_i) - x]$. Since $\lim_{i\to+\infty} 1/\lambda_i = +\infty$, we may assume without loss of generality that $\lambda_i \in (0, 1)$ for all i. Continuity of h and $\lambda_i < 1$ imply that there exists z_i such that $h(z_i) = y_i$. The definition of λ_i and strict concavity of V imply

$$\begin{aligned} V(y_i) &= V[(1-\lambda_i)x + \lambda_i h(y_i)] > (1-\lambda_i)V(x) + \lambda_i V(h(y_i)) \\ &= V(x) - \lambda_i[V(h(x)) - V(h(y_i))]. \end{aligned}$$

Thus, we have $V(h(x)) - V(h(y_i)) > [V(x) - V(y_i)]/\lambda_i$ and $h(x) - h(y_i) = (x - y_i)/\lambda_i$. From these properties and (7), we obtain

$$\begin{aligned} &V(h(x)) - V(h(y_i)) + q_x[h(y_i) - h(x)] \\ &\quad = V(h(x)) - V(h(y_i)) + p_x[h(y_i) - h(x)] + (q_x - p_x)[h(y_i) - h(x)] \\ &\quad > [V(x) - V(y_i) + p_x(y_i - x)]/\lambda_i + (q_x - p_x)[h(y_i) - h(x)] \\ &\quad \geqslant [V(x) - V(y_i) + p_x(y_i - x)]/\lambda_i. \end{aligned}$$

Since $\lim_{i\to+\infty} \lambda_i = 0$, we find that this inequality is a contradiction to (4). This completes the proof. ■

Theorem 1 rules out as an optimal policy function any function that has slope $+\infty$ at a fixed point x which is either in the interior of the state space (this is the case if $z = x$ in the theorem) or can be reached along a trajectory of h emanating from the interior of X. For example, the function $h(x) = x^{1/3}$ defined on $X = [-1, 1]$ cannot be an optimal policy function of any problem (Ω, U, ρ) that satisfies A1–A4. It is interesting to note that there are optimal policy functions which have slope $-\infty$ at an interior fixed point (see Mitra and Sorger, 1999, Example 1). It is well known that Theorem 1 does not hold if the interiority assumption is not made. The following example demonstrates that the theorem also fails if x is not a fixed point.

2) p_x is a subgradient because the sub-differential $\partial V(x)$ is compact.

Example 1

Let $X = [\underline{x}, 2]$, where $\underline{x}$ is any non-positive real number, and define the function h by

$$h(x) = \begin{cases} 0 & \text{if } x \leqslant 1, \\ \sqrt{x-1} & \text{if } x > 1. \end{cases}$$

Note that h is continuous but not Lipschitz-continuous because it has the slope $+\infty$ at the interior point $x = 1$. We now show that the condition stated in Proposition 1(ii) is satisfied for all $\rho \in (0, 1)$ provided that $V(x) = 32x - x^4$. To this end, first note that V is continuous, strictly concave and strictly increasing. Moreover, V is continuously differentiable on all of $\mathbb{R}$ such that the subgradients in (4) can be replaced by the usual derivatives. The term in brackets on the right-hand side of (4) is given by

$$V(x) - V(y) + V'(x)(y - x) = 3x^4 - 4x^3 y + y^4. \tag{8}$$

We have to consider four different cases.

Case 1 ($x \leqslant 1$ and $y \leqslant 1$). In this case we have $h(x) = h(y) = 0$ such that the left-hand side of (4) equals 0. Thus, (4) holds independently of ρ.

Case 2 ($x \leqslant 1$ and $y > 1$). In this case we have $h(x) = 0$ and $h(y) = \sqrt{y-1} \leqslant 1$. This yields $V(h(x)) - V(h(y)) + V'(h(x))[h(y) - h(x)] = (y-1)^2$. Together with (8), this shows that (4) is satisfied provided that $\rho \leqslant (3x^4 - 4x^3 y + y^4)/(y-1)^2$. We claim that the right-hand side of this inequality is greater than or equal to 1, such that the inequality holds for all $\rho \in (0, 1)$. The claim is true if $f(x, y) = 3x^4 - 4x^3 y + y^4 - (y-1)^2 \geqslant 0$ holds for all $x \leqslant 1$ and $y > 1$. Since $f_x(x, y) = 12x^2(x - y) < 0$, it follows that $f(x, y) \geqslant f(1, y) = (1-y)^2(2 + 2y + y^2) > 0$, and the claim is proved.

Case 3 ($x > 1$ and $y \leqslant 1$). In this case we have $h(x) = \sqrt{x-1} \leqslant 1$ and $h(y) = 0$. This yields $V(h(x)) - V(h(y)) + V'(h(x))[h(y) - h(x)] = 3(x-1)^2$. Together with (8), this shows that (4) is satisfied provided that $\rho \leqslant (3x^4 - 4x^3 y + y^4)/[3(x-1)^2]$. As in case 2, it is sufficient to verify that the right-hand side of this inequality is greater than or equal to 1. This is true if $f(x, y) = 3x^4 - 4x^3 y + y^4 - 3(x-1)^2 \geqslant 0$ holds for all $x > 1$ and $y \leqslant 1$. Since $f_y(x, y) = 4(y^3 - x^3) < 0$, it follows that $f(x, y) \geqslant f(x, 1) = (1-x)^2(-2 + 2x + 3x^2) > 0$, and the claim is proved.

Case 4 ($x > 1$ and $y > 1$). In this case we have $h(x) = a = \sqrt{x-1}$ and $h(y) = b = \sqrt{y-1}$. Note that $a \in (0, 1]$ and $b \in (0, 1]$ and that $x = a^2 + 1$ and $y = b^2 + 1$. Substituting this into (8) yields

$$V(x) - V(y) + V'(x)(y - x) = (a-b)^2(a+b)^2(6 + 8a^2 + 3a^4 + 4b^2 + 2a^2b^2 + b^4).$$

Moreover,

$$V(h(x)) - V(h(y)) + V'(h(x))[h(y) - h(x)] = (a-b)^2(3a^2 + 2ab + b^2).$$

Thus, (4) is satisfied provided that

$$\rho \leqslant (a+b)^2(6 + 8a^2 + 3a^4 + 4b^2 + 2a^2b^2 + b^4)/(3a^2 + 2ab + b^2).$$

Again, we show that the right-hand side of this inequality is greater than or equal to 1. This is equivalent to

$$f(a, b) = (a+b)^2(6 + 8a^2 + 3a^4 + 4b^2 + 2a^2b^2 + b^4) - (3a^2 + 2ab + b^2) \geqslant 0$$

for all $a \in (0, 1]$ and $b \in (0, 1]$. A simple calculation shows that

$$f(a, b) = 3a^2 + 10ab + 5b^2 + (a+b)^2(8a^2 + 3a^4 + 4b^2 + 2a^2b^2 + b^4).$$

Thus, $f(a, b)$ is a sum of positive terms, and the claim is proved.

These results prove that h can be the optimal policy function of a problem satisfying A1–A4 and A6 for any discount factor $\rho \in (0, 1)$. ∎

4. Robustness of topological chaos under mild discounting

In this section we prove that topological chaos is a robust phenomenon in dynamic optimization problems satisfying the standard continuity and convexity assumptions even for discount factors arbitrarily close to 1. We proceed in several steps.

Theorem 2. *For every discount factor $\rho \in (0, 1)$ there exists an optimization model (Ω, U, ρ) satisfying A1–A6 that has an optimal policy function with positive topological entropy.*

Proof. Consider the family of functions $h_\mu : [0, 1] \mapsto [0, 1]$ defined by

$$h_\mu(x) = \begin{cases} \mu x & \text{for } x \in [0, \frac{1}{2}], \\ \mu(1-x) & \text{for } x \in (\frac{1}{2}, 1], \end{cases}$$

where μ is a parameter taking values in the interval (1, 2]. For each $\mu \in (1, 2]$, the function h_μ has a tent-shaped graph and it is Lipschitz-continuous with Lipschitz constant μ. Thus, Proposition 2 shows that, for any discount factor $\rho_\mu \in (0, 1/\mu^2]$, there exists a dynamic optimization model $(\Omega_\mu, U_\mu, \rho_\mu)$ such that h_μ is the optimal policy function of this model. Since h_μ has constant slope μ, it follows from Alseda *et al.* (1993, Corollary 4.3.13) that the topological entropy of h_μ is $\ln \mu > 0$. Combining these results, and letting μ converge to 1, the theorem is proved. ∎

Remark 2. In this paper we use the term "h exhibits topological chaos" as synonymous with "h has positive topological entropy". There is complete agreement about what is meant by the latter expression (see e.g. Alseda *et al.*, 1993, pp. 188–190). While there is some disagreement about what is meant by the former expression, one prevalent use of the term "topological chaos" is as follows. A continuous function $h : X \mapsto X$ is said to exhibit topological chaos if h has a periodic point whose period is not a power of 2. The principal result connecting this concept of topological chaos and positive topological entropy is that *a continuous function $h : X \mapsto X$ exhibits topological chaos if and only if it has positive topological entropy.* (See Block and Coppel, 1992, p. 218, where functions exhibiting topological chaos in the above sense are referred to simply as "chaotic".) This result justifies our use of the two terms interchangeably.

It should be noted that the proof of Proposition 2 (see Mitra and Sorger, 1999) uses the transition possibility set $\Omega = X \times X$. This implies that, in Theorem 2 above, we may also assume that $\Omega = X \times X$. We maintain this assumption for the rest of the section.

For every positive integer n, let $U_n : \Omega \mapsto \mathbb{R}$ be a utility function and let ρ_n be a discount factor. Assume that $\lim_{n\to+\infty} \rho_n = \rho$ and that U_n converges uniformly to U. Moreover, assume that the problems (Ω, U, ρ) and (Ω, U_n, ρ_n), $n \in \{1, 2, \ldots\}$, satisfy A1–A4. Denote by V_n and h_n the optimal value function and the optimal policy function, respectively, of (Ω, U_n, ρ_n). Analogously, denote by V and h the optimal value function and the optimal policy function, respectively, of (Ω, U, ρ).

Lemma 1. *Under the assumptions mentioned above, it holds that V_n converges uniformly to V and that h_n converges uniformly to h.*

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

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

$$\frac{M}{1-\rho}\left(\frac{1+\rho}{2}\right)^{T+1} \leqslant \frac{\epsilon}{4}.$$

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

$$\sup\{|U_n(x, y) - U(x, y)| \,|\, (x, y) \in \Omega\} < \epsilon(1-\rho)/8,$$

$$\rho_n \leqslant (1+\rho)/2,$$

$$\max\{|\rho_n^t - \rho^t| \,|\, t \in \{1, 2, \ldots, T\}\} \leqslant \epsilon/(4MT).$$

Consider an arbitrary state $x \in X$ and fix an arbitrary $n \geqslant N$. Let $(x_t)_{t=0}^{+\infty}$ be an optimal path from initial state x for (Ω, U, ρ). Note that this path is also feasible for the model (Ω, U_n, ρ_n). The following chain of inequalities holds because of the above properties:

$$\begin{aligned}
V_n(x) &\geqslant \sum_{t=0}^{+\infty} \rho_n^t U_n(x_t, x_{t+1}) \\
&\geqslant \sum_{t=0}^{T} \rho_n^t U_n(x_t, x_{t+1}) \\
&= \sum_{t=0}^{T} (\rho_n^t - \rho^t) U_n(x_t, x_{t+1}) + \sum_{t=0}^{T} \rho^t U_n(x_t, x_{t+1}) \\
&\geqslant -(\epsilon/4) + \sum_{t=0}^{T} \rho^t U_n(x_t, x_{t+1}) \\
&\geqslant -(\epsilon/4) + \sum_{t=0}^{T} \rho^t U(x_t, x_{t+1}) - (\epsilon/8) \\
&> \sum_{t=0}^{+\infty} \rho^t U(x_t, x_{t+1}) - \sum_{t=T+1}^{+\infty} \rho^t U(x_t, x_{t+1}) - (\epsilon/2) \\
&\geqslant V(x) - (\epsilon/4) - (\epsilon/2) \\
&> V(x) - \epsilon.
\end{aligned}$$

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

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

$$V(x_0) \geqslant U(x_0, z_0) + \rho V(z_0) + \epsilon.$$

Because of uniform convergence of U_n to U and V_n to V, convergence of ρ_n to ρ, as well as continuity of U and V, one can find an integer N such that, for all $n \geqslant N$ and all $x \in X$, the following properties hold:

$$|V_n(x) - V(x)| < \epsilon/8,$$

$$|V(x_n) - V(x_0)| < \epsilon/8,$$

$$|U(x_n, h_n(x_n)) - U(x_0, z_0)| < \epsilon/8,$$

$$|V(h_n(x_n)) - V(z_0)| < \epsilon/8,$$

$$|\rho_n - \rho| < \epsilon(1-\rho)/(8M).$$

From these conditions, we obtain

$$\begin{aligned} V_n(x_n) &= [V_n(x_n) - V(x_n)] + [V(x_n) - V(x_0)] + V(x_0) \\ &\geqslant -(\epsilon/8) - (\epsilon/8) + V(x_0) \\ &\geqslant U(x_0, z_0) + \rho V(z_0) + (3\epsilon/4) \\ &\geqslant U(x_n, h_n(x_n)) + \rho V(h_n(x_n)) + (\epsilon/2) \\ &= U(x_n, h_n(x_n)) + \rho_n V(h_n(x_n)) + (\epsilon/2) + (\rho - \rho_n)V(h_n(x_n)) \\ &\geqslant U(x_n, h_n(x_n)) + \rho_n V(h_n(x_n)) + (3\epsilon/8) \\ &\geqslant U(x_n, h_n(x_n)) + \rho_n V_n(h_n(x_n)) + (\epsilon/4) \\ &= V_n(x_n) + (\epsilon/4). \end{aligned}$$

Clearly, this is a contradiction, and the result is proved. ■

We are now prepared to prove the second main result of this section. Denote the space of all dynamic optimization models (Ω, U, ρ) satisfying assumptions A1–A6 and $\Omega = X \times X$ by $\mathscr{M}_X$ and define the metric

$$\delta((\Omega, U, \rho), (\Omega, U', \rho')) = \max\{|U(x, y) - U'(x, y)| | (x, y) \in \Omega\} + |\rho - \rho'|$$

on this space.

Theorem 3. *Assume that* $(\Omega, U, \rho) \in \mathscr{M}_X$ *has an optimal policy function with positive topological entropy. Then there exists* $\epsilon > 0$ *such that all models* $(\Omega, U', \rho') \in \mathscr{M}_X$ *with* $\delta((\Omega, U, \rho), (\Omega, U', \rho')) < \epsilon$ *have optimal policy functions with positive topological entropy.*

Proof. The result follows at once from Lemma 1 and from the fact that the topological entropy is a lower semi-continuous function on the space of all continuous functions of X into itself endowed with the topology of uniform convergence (see Alseda *et al.*, 1993, Theorem 4.5.2). ■

Theorems 2 and 3 together prove our claim that topological chaos is a robust phenomenon in dynamic optimization models satisfying A1–A6 even under arbitrary mild discounting.

5. Exact discount factor restrictions

In this section we demonstrate how the results stated in Section 3 can be used to characterize the set of discount factors that are compatible with the optimality of a given continuous mapping $h: X \to X$. We do this by considering two important examples: the logistic map $h(x) = 4x(1-x)$, and the tent map $h(x) = 1 - |2x - 1|$, both of which are defined on the set $X = [0, 1]$. These two maps are among the best known examples of chaotic maps. (For previous studies of the rationalizability of these functions, we refer to the literature mentioned in the Introduction.)

Theorem 4. *Let $X = [0, 1]$ be the state space and $\rho \in (0, 1)$ the discount factor, and define $h : X \mapsto X$ by $h(x) = 4x(1 - x)$. The following two conditions are equivalent:*

(i) There exists a transition possibility set Ω and a utility function U such that the model (Ω, U, ρ) satisfies A1–A6 and such that h is the optimal policy function of (Ω, U, ρ).
(ii) The discount factor satisfies $\rho \in (0, 1/16]$.

Proof. (*a*) Since h is Lipschitz-continuous with Lipschitz constant $L = 4$, it follows from Proposition 2 that for any $\rho \leqslant 1/16$ one can find Ω and U such that (Ω, U, ρ) satisfies A1–A6 and has h as its optimal policy function.

(*b*) Now assume that (Ω, U, ρ) is given such that A1–A6 are satisfied and (h, V) is the solution of this model. We have to show that $\rho \leqslant 1/16$. Assume to the contrary that $\rho > 1/16$; then one can find numbers $\lambda \in (0, 1)$ and $\theta \in (1, \infty)$ such that

$$16\lambda^2\rho > \theta. \tag{9}$$

From A5, it follows that there exist positive numbers α and β such that the optimal value function V is α-concave and $(-\beta)$-convex. Let us choose a sufficiently large integer T such that

$$(\beta/\alpha)^{1/(T+1)} \leqslant \theta. \tag{10}$$

Because h is continuously differentiable and $h'(0) = 4$, we can find $\epsilon > 0$ such that

$$h'(\epsilon) \geqslant 4\lambda. \tag{11}$$

Now let us define $x_0 = \epsilon/4^{T+1}$ and $x_{t+1} = h(x_t)$ for $t \in \{0, 1, \ldots, T\}$. Since $h'(x) \in (4\lambda, 4)$ for all $x \in (0, \epsilon)$, it follows that $0 < x_0 < x_1 < \ldots < x_{T+1} < \epsilon$. From Proposition 1(i), it follows that, for every $t \in \{0, 1, \ldots, T + 1\}$, there exists $p_t \in \partial V(x_t)$ such that, for all $y \in X$ and all $t \in \{0, 1, \ldots, T\}$,

$$V(h(x_t)) - V(h(y)) + p_{t+1}[h(y) - h(x_t)] \leqslant (1/\rho)[V(x_t) - V(y) + p_t(y - x_t)].$$

Combining these $T + 1$ inequalities and choosing $y = 0$ yields

$$V(x_{T+1}) - V(0) + p_{T+1}(0 - x_{T+1}) \leqslant (1/\rho^{T+1})[V(x_0) - V(0) + p_0(0 - x_0)].$$

Since V is α-concave and $(-\beta)$-convex, we obtain from this inequality

$$\alpha x_{T+1}^2 \geqslant (\beta/\rho^{T+1})x_0^2. \tag{12}$$

Since h is concave, $h(0) = 0$, and $x_t < \epsilon$ for all $t \in \{0, 1, \ldots, T\}$, we get $x_{t+1} = h(x_t) > h'(x_t)x_t > h'(\epsilon)x_t$. Thus, $x_{T+1} > [h'(\epsilon)]^{T+1}x_0 \geqslant (4\lambda)^{T+1}x_0$, by (11). This inequality can also be written as $x_{T+1}^2 \geqslant (16\lambda^2)^{T+1}x_0^2$. Together with (12), this shows that $\rho^{T+1}(16\lambda^2)^{T+1} \leqslant \beta/\alpha$. Using (10), it follows that $16\lambda^2\rho \leqslant \theta$, which is a contradiction to (9). This completes the proof. ∎

The above theorem shows that the logistic map can be an optimal policy function of a regular optimization problem if and only if the discount factor does not exceed 1/16. We do not know if the logistic map can be rationalized by an optimization model with a higher discount factor if the regularity assumption A5 is not required.

The second example that we consider in this section is the tent map $h(x) = 1 - |2x - 1|$. Although this function has properties very similar to the logistic map, it can be rationalized for a larger set of discount factors.[3]

Theorem 5. *Let* $X = [0, 1]$ *be the state space and* $\rho \in (0, 1)$ *the discount factor, and define* $h : X \mapsto X$ *by* $h(x) = 1 - |2x - 1|$. *The following two conditions are equivalent:*

(i) There exists a transition possibility set Ω *and a utility function* U *such that the model* (Ω, U, ρ) *satisfies A1–A4 and such that* h *is the optimal policy function of* (Ω, U, ρ).
(ii) The discount factor satisfies $\rho \in (0, 1/4]$.

Proof. (*a*) Since h is Lipschitz-continuous with Lipschitz constant $L = 2$, it follows from Proposition 2 that, for any $\rho \leqslant 1/4$, one can find Ω and U such that (Ω, U, ρ) satisfies A1–A4 and rationalizes h.

(*b*) Now assume that (Ω, U, ρ) is given such that A1–A4 are satisfied and such that (h, V) is the solution of this model. We have to show that $\rho \leqslant 1/4$. To this end we proceed in four steps. In the first one we show that the optimal value function must have finite one-sided derivatives at all states, in the second and third steps we derive two different discount factor restrictions, and in the final step we combine the two restrictions to arrive at the one stated in the theorem.

Step 1. Because $\frac{1}{2}$ is in the interior of X, it follows that V is sub-differentiable at $\frac{1}{2}$. From Proposition 1(i), it follows that V is also sub-differentiable at $h(\frac{1}{2}) = 1$. The left-hand derivative $V'_{-}(1)$ cannot be $-\infty$ because the direction -1 points into X at 1, and it cannot be $+\infty$ because that would violate sub-differentiability of V at 1. Because V is sub-differentiable at 1, it follows from Proposition 1(i) that V is sub-differentiable at $h(1) = 0$. By a similar argument as above, the right-hand derivative $V'_{+}(0)$ can be neither $-\infty$ nor $+\infty$. From these properties and the concavity of V, we obtain $+\infty > V'_{+}(0) \geqslant V'_{-}(x) \geqslant V'_{+}(x) \geqslant V'_{-}(1) > -\infty$ for all $x \in (0, 1)$.

Step 2. Let x and y be arbitrary elements of (0, 1). From Proposition 1(i), it follows that there exist subgradients $p_x \in V(x)$, $q_x \in V(h(x))$, $p_y \in V(y)$ and $q_y \in V(h(y))$ such that (5) holds. Choose $\epsilon \in (0, \frac{1}{4})$ and let $x = \frac{1}{2}$ and $y = 1 - \epsilon$. Then $h(x) = 1$ and $h(y) = 2\epsilon$. Thus, there are $p_x \in \partial V(\frac{1}{2})$, $q_x \in \partial V(1)$, $p_y \in \partial V(1 - \epsilon)$ and $q_y \in \partial V(2\epsilon)$ such that

$$(q_y - q_x)(1 - 2\epsilon) \leqslant (1/\rho)(p_x - p_y)[(\tfrac{1}{2}) - \epsilon].$$

From our choice of x and y, we get $p_x \leqslant V'_{-}(\frac{1}{2})$, $p_y \geqslant V'_{+}(1 - \epsilon)$, $q_x \leqslant V'_{-}(1)$ and $q_y \geqslant V'_{+}(2\epsilon)$. Using these facts, we obtain from the above inequality

$$\rho \leqslant \frac{[(\frac{1}{2}) - \epsilon][V'_{-}(\frac{1}{2}) - V'_{+}(1 - \epsilon)]}{(1 - 2\epsilon)[V'_{+}(2\epsilon) - V'_{-}(1)]}.$$

3) Similarity of the logistic map and the tent map derives from their topological equivalence. Two maps $h_1 : X_1 \mapsto X_1$ and $h_2 : X_2 \mapsto X_2$ are called topologically equivalent if there exists a homeomorphism $f : X_1 \mapsto X_2$ such that $f(h_1(x)) = h_2(f(x))$ for all $x \in X_1$.

In the limit, as ϵ approaches 0 this yields, because of the continuity properties of one-sided derivatives, that

$$\rho \leqslant \frac{V'_-(\frac{1}{2}) - V'_-(1)}{2[V'_+(0) - V'_-(1)]}. \tag{13}$$

Step 3. Choose $\epsilon \in (0, \frac{1}{4})$ and let $x = \epsilon$ and $y = (\frac{1}{2}) - \epsilon$. Then $h(x) = 2\epsilon$ and $h(y) = 1 - 2\epsilon$. As in step 2, it follows from Proposition 1(i) and (5) that there exist $p_x \in \partial V(\epsilon)$, $q_x \in \partial V(2\epsilon)$, $p_y \in \partial V((\frac{1}{2}) - \epsilon)$ and $q_y \in \partial V(1 - 2\epsilon)$ such that

$$(q_x - q_y)(1 - 4\epsilon) \leqslant (1/\rho)(p_x - p_y)[(\tfrac{1}{2}) - 2\epsilon].$$

Moreover, we have $p_x \leqslant V'_-(\epsilon)$, $p_y \geqslant V'_+((\frac{1}{2}) - \epsilon)$, $q_x \geqslant V'_+(2\epsilon)$ and $q_y \leqslant V'_-(1 - 2\epsilon)$. Together with the above inequality, this yields

$$\rho \leqslant \frac{[(\frac{1}{2}) - 2\epsilon][V'_-(\epsilon) - V'_+((\frac{1}{2}) - \epsilon)]}{(1 - 4\epsilon)[V'_+(2\epsilon) - V'_-(1 - 2\epsilon)]}.$$

As in step 2, we take the limit as ϵ approaches 0, which yields

$$\rho \leqslant \frac{V'_+(0) - V'_-(\frac{1}{2})}{2[V'_+(0) - V'_-(1)]}. \tag{14}$$

Step 4. Define $\lambda = [V'_+(0) - V'_-(\frac{1}{2})]/[V'_+(0) - V'_-(1)]$. Then we can write (13) and (14) as $\rho \leqslant (1 - \lambda)/2$ and $\rho \leqslant \lambda/2$, respectively. Thus, we must have $\rho \leqslant (\frac{1}{2}) \min\{\lambda, 1 - \lambda\} \leqslant \frac{1}{4}$. This completes the proof. ■

Theorem 5 provides an exact discount factor restriction for the rationalizability of the tent map. Note that, in contrast to Theorem 4, the regularity assumption A5 is not needed for this result. It is clear from the proof, however, that, whenever $\rho \leqslant \frac{1}{4}$, one can find a model (Ω, U, ρ) that rationalizes the tent map and satisfies not only A1–A4, but also A5 and A6.

Final version accepted 14 December 1998.

REFERENCES

Alseda, L., J. Llibre and M. Misiurewicz (1993) *Combinatorial Dynamics and Entropy in Dimension One*, Singapore: World Scientific.

Block, L. S. and W. A. Coppel (1992) *Dynamics in One Dimension*, Berlin: Springer-Verlag.

Boldrin, M. and L. Montrucchio (1986) "On the Indeterminacy of Capital Accumulation Paths", *Journal of Economic Theory*, Vol 40, pp. 26–39.

Deneckere, R. and S. Pelikan (1986) "Competitive Chaos", *Journal of Economic Theory*, Vol. 40, pp. 13–25.

Hewage, T. U. and D. A. Neumann (1990) "Functions Not Realizable as Policy Functions in an Optimal Growth Model", Discussion Paper, Bowling Green State University.

McKenzie, L. W. (1986) "Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics", in K. Arrow and M. Intriligator (eds.), *Handbook of Mathematical Economics*, Vol. 3, Amsterdam: North-Holland, pp. 1281–1355.

Mitra, T. (1996) "On the Nature of Policy Functions of Dynamic Optimization Models", Working Paper, Cornell University.

—— and G. Sorger (1999) "Rationalizing Policy Functions by Dynamic Optimization", *Econometrica*, Vol. 67, pp. 375–392.

Montrucchio, L. (1994) “Dynamic Complexity of Optimal Paths and Discount Factors for Strongly Concave Problems”, *Journal of Optimization Theory and Applications*, Vol. 80, pp. 385–406.
Nishimura, K. and M. Yano (1995) “Non-linear Dynamics and Chaos in Optimal Growth: an Example”, *Econometrica*, Vol. 63, pp. 981–1001.
—— G. Sorger and M. Yano (1994) “Ergodic Chaos in Optimal Growth Models with Low Discount Rates”, *Economic Theory*, Vol. 4, pp. 705–717.
—— T. Shigoka and M. Yano (1998) “Interior Optimal Chaos with Arbitrarily Low Discount Rates”, *Japanese Economic Review*, Vol. 49, pp. 223–233.
Royden, H. L. (1988) *Real Analysis*, New York: Macmillan.
Sorger, G. (1992a) “On the Minimum Rate of Impatience for Complicated Optimal Growth Paths”, *Journal of Economic Theory*, Vol. 56, pp. 160–179.
—— (1992b) *Minimum Impatience Theorems for Recursive Economic Models*, Heidelberg: Springer-Verlag.
—— (1994) “Policy Functions of Strictly Concave Optimal Growth Models”, *Ricerche Economiche*, Vol. 48, pp. 195–212.
—— (1995) “On the Sensitivity of Optimal Growth Paths”, *Journal of Mathematical Economics*, Vol. 24, pp. 353–369.
Stokey, N. L. and R. E. Lucas, Jr (1989) *Recursive Methods in Economic Dynamics*, Cambridge, MA: Harvard University Press.