13

On the Relationship between Discounting and Entropy of Dynamic Optimization Models

Tapan Mitra

1. Introduction

In a standard aggregative dynamic optimization framework (Ω, u, δ), where Ω is the transition possibility (technology) set, u is a (reduced form) utility function defined on this set, and $0 < \delta < 1$ is a discount factor, the relation between the magnitude of the discount factor and the extent of complicated behavior generated by the corresponding (optimal) policy function has been a topic of extensive study.

Boldrin and Montrucchio (1986) showed that any twice continuously differentiable function can be a policy function of an appropriately defined dynamic optimization model. However, when the C^2 function was taken to be the logistic map ($h(x) = 4x(1-x)$ for $x \in [0, 1]$), which is well-known to exhibit complicated behavior, the dynamic optimization model (obtained by using their constructive method) for which the logistic map was the policy function was seen to have an extremely small discount factor (about 0.01). Subsequently, Sorger (1992a) showed that if the policy function of *any* dynamic optimization model is the logistic map, then its associated discount factor must be smaller than 0.5.

These results suggested that substantial discounting might be *necessary* to obtain complicated (or "chaotic") optimal behavior. This was confirmed by Sorger (1992b) when he showed, using the theory of stochastic dominance, that if any dynamic optimization model (Ω, u, δ) exhibits a period-three cycle, then the discount

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

factor, δ, must satisfy

$$\delta < (\sqrt{5} - 1)/2 \approx 0.618 \tag{1.1}$$

In subsequent work, Sorger (1994) refined the above bound to

$$\delta < 0.5479 \tag{1.2}$$

The bound was further refined to

$$\delta < [(\sqrt{5} - 1/2]^2 \approx 0.3819 \tag{1.3}$$

in Mitra (Chapter 11) and Nishimura and Yano (Chapter 12). Furthermore, the bound in (1.3) was shown to be "exact" in the sense that whenever $\delta < [(\sqrt{5} - 1)/2]^2$, one can construct a transition possibility set, Ω, and a reduced form utility function, u, such that the dynamic optimization model (Ω, u, δ) has an optimal program exhibiting a period-three cycle.

A period-three cycle is a special case of what is known as "topological chaos", which occurs whenever there is a periodic cycle of a period not equal to a power of 2. [Topological chaos implies the existence of an infinite number of periodic cycles of distinct periodicities, as well as the existence of an uncountable "scrambled set"; see Section 2 for details]. Thus, if we focus on the existence of any periodic point of period $q = np$, where $n > 1$ is an odd integer and $p = 2^k$, with k a non-negative integer, it can be shown [see Mitra (Chapter 11)] that, combining (1.3) with the "Sarkovskii order" (see Proposition 2.1 below), the following bound on the discount factor can be obtained:

$$\delta < [(\sqrt{5} - 1)/2]^{1/2^k} \tag{1.4}$$

In studying the closely related phenomenon of "turbulence" (see Section 2 for a definition), Mitra (1994) established that if any dynamic optimization model (Ω, u, δ) exhibits turbulence, then the discount factor, δ, must satisfy

$$\delta < (1/4) \tag{1.5}$$

Furthermore, the bound in (1.5) is "exact" in the sense that whenever $\delta < 0.25$, one can construct a transition possibility set, Ω, and a utility function, u, such that the dynamic optimization model (Ω, u, δ) exhibits turbulence. Using the "Sarkovskii order" and (1.5), the bound obtained in (1.4) can be refined to

$$\delta < [1/2]^{1/2^k} \tag{1.6}$$

as shown in Mitra (1994).

All these results certainly indicate a close relationship between discounting and complex optimal behavior. However, this relationship could be made more precise if we had a convenient numerical measure of complicated behavior. From the topological point of view, such a numerical measure is the "topological entropy" of a dynamical system. Thus, one could proceed to examine whether the upper bound on the discount factor goes on decreasing as the extent of complicated optimal behavior goes on increasing. This approach was pioneered by Montrucchio (1994), who established (under a strong concavity assumption on the utility function) that

$$\delta \leq 1/e^{\psi(h,A)} \tag{1.7}$$

where A is a compact, invariant set contained in the interior of the "state space" of the dynamic optimization model, and $\psi(h, A)$ is the topological entropy on the set A of the policy function, h, when the discount factor is δ. Montrucchio and Sorger (1994) have shown that the inequality holds even when the strong concavity assumption on the utility function is replaced by a relatively mild strict concavity assumption on the utility function [of the type used in Section 3 of this chapter].

In this chapter, we do three things. First, we establish the Montrucchio-Sorger result (1.7) by using what I have called the "value-loss approach" to minimum impatience problems. Focusing on "value-losses" (that one suffers by deviating from certain price-supported activities) has been the cornerstone of the approach of McKenzie [see McKenzie (1986) for a comprehensive survey] to problems of "turnpike" theory. It appears to be a natural concept to focus on for the study of complicated optimal behavior as well, and the inequality (1.7) is seen to rest firmly on a version of the well-known "value-loss lemma", which was introduced to (final-state) turnpike theory by Radner (1961), and which has occupied center-stage in this [and in consumption-oriented turnpike theory] literature thereafter.

Second, under an additional assumption of "bounded steepness" of the utility function (see Section 5 for details) we extend the Montrucchio-Sorger result to establish that

$$\delta \leq 1/e^{\psi(h,X)} \tag{1.8}$$

where X is the compact state space itself. A problem in applying the Montrucchio-Sorger result to obtain discount factor restrictions for topological chaos is that, in general, given a policy function it is difficult to identify compact invariant sets in the *interior* of the state space. There is, of course, no such difficulty in applying formula (1.8).

Noting that topological chaos may not be "observable", we consider a natural alternate measure of chaos, metric entropy, and show (using a mathematical result comparing the magnitude of the topological entropy with the metric entropy of a dynamical system) that

$$\delta \leq 1/e^{\phi_\mu(h,X)} \tag{1.9}$$

where μ is a probability measure on the Borel σ-field of X, invariant under h, and $\phi_\mu(h,X)$ the metric entropy.

Results (1.8) and (1.9) [Theorems 5.2 and 5.3 of this chapter] summarize the relationship between discounting and complicated behavior in dynamic optimization models, where "complicated behavior" is measured in the topological sense in (1.8) and in the measure-theoretic sense in (1.9).

Third, I study the implications of result (1.8) for periodic and turbulent optimal programs. The concept of topological entropy has been thoroughly studied in the theory of dynamical systems, and powerful methods have been developed for computing the topological entropy of dynamical systems, exhibiting periodic and turbulent behavior, by Block, Guckenheimer, Misiurewicz and

Young (1980). We use these results to show that the inequality (1.8) yields the discount factor restriction

$$\delta \leq (\sqrt{5} - 1)/2 \tag{1.10}$$

for period-three cycles, and

$$\delta \leq (1/2) \tag{1.11}$$

for turbulence. These are fairly strong bounds, but turn out to be the square-roots of the "exact" bounds obtained in (1.3) and (1.5) respectively. This naturally prompts us to ask whether there is a more refined relationship than (1.8) between the discount factor and the topological entropy, which would yield the exact discount factor restrictions for period-three cycles and turbulence as special cases. We leave this as an open question.

2. Chaos

2.1 Periodic Orbits

Let I be a compact interval in $\Re$, the set of reals. Let $f : I \to I$ be a continuous map of the interval I into itself. The pair (I, f) is called a *dynamical system*; I is called the *state space* and f the *law of motion* of the dynamical system.

We write $f^0(x) = x$ and for any integer $n \geq 1$, $f^n(x) = f[f^{n-1}(x)]$. If $x \in I$, the sequence $\tau(x) = \{f^n(x)\}_0^\infty$ is called the *trajectory* from (the initial condition) x. The *orbit* from x is the set $\gamma(x) = \{y : y = f^n(x) \text{ for some } n \geq 0\}$.

A point $x \in I$ is a *fixed point* of f if $f(x) = x$. A point $x \in I$ is called a *periodic point* of f if there is $k \geq 1$ such that $f^k(x) = x$. The smallest such k is called the *period* of x. [In particular, if $x \in I$ is a fixed point of f, it is periodic with period 1]. If $x \in I$ is a periodic point with period k, we also say that the orbit of x (or trajectory from x) is periodic with period k.

The following fundamental result on the existence of periodic orbits is due to Sarkovskii (1964). A good discussion of this result is contained in Block and Coppel (1992).

Proposition 2.1: *Let the positive integers be totally ordered in the following way:*

$$3 \prec 5 \prec 7 \prec 9 \prec \ldots \prec 2.3 \prec 2.5 \prec \ldots \prec 2^2.3 \prec 2^2.5$$
$$\prec \ldots \prec 2^3 \prec 2^2 \prec 2 \prec 1.$$

If f has a periodic orbit of period n and if $n \prec m$, then f also has a periodic orbit of period m.

2.2 Aperiodic Orbits

In order to study the nature of trajectories which are not periodic, it is useful to define a "scrambled" set. A set $S \subset I$ is called a *scrambled set* if it possesses the following two properties:

(i) If x, $y \in S$ with $x \neq y$, then

$$\lim_{n\to\infty} \sup |f^n(x) - f^n(y)| > 0$$

and

$$\lim_{n\to\infty} \inf |f^n(x) - f^n(y)| = 0$$

(ii) If $x \in S$ and y is any periodic point of f,

$$\lim_{n\to\infty} \sup |f^n(x) - f^n(y)| > 0$$

Thus trajectories starting from points in a scrambled set are not even "asymptotically periodic". Furthermore, for any pair of initial states in the scrambled set, the trajectories move apart and return close to each other infinitely often.

The following theorem, due to Li and Yorke (1975), is fundamental in establishing a connection between the existence of period-three cycles and the existence of an uncountable scrambled set.

Proposition 2.2: *Assume that there is some point x^* in I such that:*

$$f^3(x^*) \leq x^* < f(x^*) < f^2(x^*)$$
$$(\text{or } f^3(x^*) \geq x^* > f(x^*) > f^2(x^*)) \tag{2.1}$$

Then (i) for every positive integer $k = 1, 2, ...,$ *there is a periodic point of period* k.

(ii) there is an uncountable scrambled set $S \subset I$.

We will say that the dynamical system (I, f) exhibits *Li-Yorke chaos* if conditions (i) and (ii) of Proposition 2.2 are satisfied. It is easy to check that (I, f) exhibits Li-Yorke chaos if and only if (I, f)has a periodic point of period three.

Generalizing the Li-Yorke result, and using Sarkovskii's theorem, Li, Misiurewicz, Pianigiani, and Yorke (1982) established the following result.

Proposition 2.3: *Suppose* f *has a periodic point of period which is not a power of 2. Then (i)* f *has infinitely many periodic points of different periods; and (ii)*I *has an uncountable scrambled set.*

2.3. Turbulence[1]

The map $f : I \to I$ will be called *turbulent* if there exist compact (non-degenerate) subintervals J, K with at most one common point such that

$$J \cup K \subset f(J) \cap f(K)$$

The concept of turbulence is related to the occurrence of periodic points in the following result.

Proposition 2.4: *(i) If* $f : I \to I$ *is turbulent then* f *has periodic point of all periods.*

(ii) If $f : I \to I$ *has a periodic point of odd period* $n > 1$*, then* f^2 *is turbulent.*

[1]The definition of turbulence used here follows Block and Coppel (1992). The phenomenon is also referred to as a "2-horseshoe" by Alseda, Llibre and Misiurewicz (1993). The concept has evolved from the earlier dicussions of Ruelle and Takens (1971) and Lasota and Yorke (1977).

2.4. Topological Entropy[2]

There is a well-known measure of the extent of complicated behavior that a dynamical system can exhibit, and it is called "topological entropy". Let A be a compact subset of I, which is invariant under f (that is, $f(A) \subset A$). Let $J = (J_1, ..., J_p)$ be a finite open cover of A (that is, $\bigcup_{i=1}^{p} J_i \supset A$). Denote by $f^{-1}J$ the open cover $(f^{-1}J_1, ..., f^{-1}J_p)$. If $J = (J_1, ..., J_p)$ and $J' = (J'_1, ..., J'_q)$ are open covers of A, then their join JVJ' is the open cover consisting of all sets of the form $J_i \cap J'_j$ for all $i \in \{1, ..., p\}$, $j \in \{1, ..., q\}$. Let $N(J)$ be the minimum number of sets in the cover J which still covers A. Then, we can define

$$\psi(f, A, J) = \lim_{n \to \infty} (1/n) \log N(JVf^{-1}JV...Vf^{-n+1}J)$$

The *topological entropy* of f on A is defined as

$$\psi(f, A) = \sup_J \psi(f, A, J)$$

where the supremum is taken over all open covers J. For $A = I$, we often write $\psi(f)$ instead of $\psi(f, I)$.

We will use a more "operational" definition of topological entropy, which we now provide. A finite set $E \subset A$ is called (n, ε)-*separated* ($n = 1, 2, ...$ and $\varepsilon > 0$) if for every x, $y \in E$, $x \neq y$, there is $0 \leq k < n$ such that $|f^k(x) - f^k(y)| \geq \varepsilon$. Let $s(n, \varepsilon)$ denote the maximal cardinality of an (n, ε)-separated set. We define

$$\psi_\varepsilon(f, A) = \limsup_{n \to \infty} (1/n) \log s(n, \varepsilon)$$

and the *topological entropy* of f as

$$\Psi(f, A) = \lim_{\varepsilon \to 0} \psi_\varepsilon(f, A).$$

---

[2] The formal definition of topological entropy was given by Adler, Konheim and McAndrew (1965). Bowen (1971a) provided the more "operational" definition (which we use here). In our context, the two definitions are equivalent; for a proof, see Bowen (1971b).

The following result, due to Misiurewicz (1979), relates topological entropy to the existence of periodic points with period which is not equal to a power of 2.

Proposition 2.5: *The following two statements are equivalent: (i) f has a periodic point of period which is not a power of 2;(ii) f has positive topological entropy.*

We will say that a dynamical system (I, f) exhibits *topological chaos* if f has a periodic point whose period is not a power of 2. This definition follows Block and Coppel (1992).

A justification for adopting the above definition of topological chaos is that, while it is weaker than the period-three condition of Li and Yorke it is strong enough to ensure the existence of an uncountable scrambled set. On the other hand, the definition is weak enough so that a dynamical system (I, f) exhibits topological chaos if and only if it has positive topological entropy.

2.5. Metric Entropy

A difficulty with the notion of "topological chaos" is that it may not be "observed", and therefore might be unsuitable for the purpose of signaling "unpredictability" of outcomes of a dynamical system. If by chaos we mean unpredictability, the natural measure of it is metric entropy (also known as "measure-theoretic entropy" or "Kolmogorov-Sinai invariant"), since it measures the "uncertainty" in an experiment about the outcome, or equivalently the "information" gained by conducting the experiment.

Let $\wp$ be the Borel σ-field of I, and v a probability measure on $\wp$. Thus, $(I, \wp, v)$ is a probability space. If f is $\wp$-measurable, then v is called *invariant* under h if $v(E) = v(f^{-1}(E))$ for all E in $\wp$.

Suppose the dynamical system (I, f) has an invariant measure v. We call $J = \{J_1, ..., J_p\}$ a (finite) *measurable partition* of I if the J_i are disjoint measurable subsets of I, and $\bigcup_i J_i \supset I$. If $J = \{J_1, ..., J_p\}$ is a finite measurable partition of I, then the *entropy of the partition* J is

$$H(f, I, v, J) = -\sum_{i=1}^{p} v(J_i) \ \log \ v(J_i)$$

If $J = \{J_1, ..., J_p\}$ is a measurable partition, we denote by $f^{-1}J$ the measurable partition $\{f^{-1}J_1, ..., f^{-1}J_p\}$. If $J = \{J_1, ..., J_p\}$ and $J' = \{J'_1, ..., J'_q\}$ are measurable partitions of X, then JVJ' is the measurable partition consisting of sets of the form $J_i \cap J'_j$ for all $i \in \{1, ..., p\}$, $j \in \{1, ..., q\}$. We can then define

$$\Phi(f, I, v, J) = \lim_{n\to\infty} (1/n) H(f, I, v, JVf^{-1}JV...Vf^{-n+1}J)$$

Finally, the *metric entropy* is defined as

$$\Phi_v(f, I) = \sup_J \Phi(f, I, v, J)$$

where the supremum is taken over all partitions J of I for which $\Phi(f, I, v, J) < \infty$.

3. Dynamic Optimization

3.1 The Model

The framework is described by a triplet (Ω, u, δ), where Ω, a subset of $\Re_+ \times \Re_+$, is a *transition possibility set*, $u : \Omega \to \Re$ is a *utility function* defined on this set, and δ is the *discount factor* satisfying $0 < \delta < 1$.

The transition possibility set describes the states $z \in \Re_+$ that it is possible to go to tomorrow, if one is in state $x \in \Re_+$ today. We define a correspondence $\Gamma : \Re_+ \to \Re_+$ by $\Gamma(x) = \{y \in \Re_+ : (x, y) \in \Omega\}$ for each $x \in \Re_+$.

A program $\{x_t\}_0^\infty$ from $\mathbf{x} \in \Re_+$ is a sequence satisfying

$$x_0 = \mathbf{x} \text{ and } (x_t, x_{t+1}) \in \Omega \text{ for } t \geq 0$$

If one is in state x today and one moves to state z tomorrow (with $(x, z) \in \Omega$) then there is an immediate utility (or “reward” or “return”) generated, measured by the utility function, u.

The discount factor, δ, is the weight assigned to tomorrow's utility (compared to today's) in the objective function. The *discount rate* (associated with the discount factor, δ) is given by $\rho = (1/\delta) - 1$.

The following assumptions are imposed on the transition possibility set, Ω, and the utility function, u:

A1: *(i)* $(0,0) \in \Omega$, *(ii)* $(0,z) \in \Omega$ *implies* $z = 0$.
A2: Ω *is (i) closed, and (ii) convex.*
A3: *There is* $\xi > 0$ *such that* $(x,z) \in \Omega$ *and* $x \geq \xi$ *implies* $z < x$.
A4: *If* $(x,z) \in \Omega$ *and* $x' \geq x$, $0 \leq z' \leq z$ *then* $(x',z') \in \Omega$.

Clearly, we can pick $0 < \zeta < \xi$, such that if $x > \zeta$ and $(x,z) \in \Omega$, then $z < x$. It is straightforward to verify that if $(x,z) \in \Omega$, then $z \leq \max(\zeta, x)$. It follows from this that if $\{x_t\}_0^\infty$ is a program from $\mathbf{x} \in \Re_+$, then $x_t \leq \max(\zeta, \mathbf{x})$ for $t \geq 0$. In particular, if $\mathbf{x} \leq \zeta$, then $x_t \leq \zeta$ for $t \geq 0$. This leads us to choose the closed interval, $[0,\zeta]$ as the natural *state space* of our model, which we will denote by X. We denote the interval $[0,\xi]$ by Y.

The following assumptions are imposed on the utility function, u:

A5: u *is concave on* Ω*; further if* (x,z) *and* (x',z') *are in* Ω*, and* $x \neq x'$*, then for every* $0 < \lambda < 1$,
$u(\lambda(x,z)+$
$(1-\lambda)(x',z')) > \lambda u(x,z) + (1-\lambda)u(x',z')$.
A6: u *is upper semi-continuous on* Ω.
A7: *If* $x, x' \in Y$, $(x,z) \in \Omega$, $x' \geq x$ *and* $0 \leq z' \leq z$, *then* $u(x',z') \geq u(x,z)$.

We will refer to a triplet (Ω, u, δ) satisfying (A.1) - (A.7) as a *dynamic optimization model.*

A program $\{\hat{x}_t\}_0^\infty$ from $\mathbf{x} \geq 0$ is an *optimal program* if

$$\sum_0^\infty \delta^t u(x_t, x_{t+1}) \leq \sum_0^\infty \delta^t u(\hat{x}_t, \hat{x}_{t+1})$$

for every program $\{x_t\}_0^\infty$ from $\mathbf{x}$.

Under (A1)-(A7), there is a unique optimal program from every $\mathbf{x} \in \Re_+$.

3.2 Value and Policy Functions

The *value function* $V : \Re_+ \to \Re$ is defined by

$$V(x) = \sum_0^\infty \delta^t u(\hat{x}_t, \hat{x}_{t+1})$$

where $\{\hat{x}_t\}_0^\infty$ is the optimal program from $x \in \Re_+$.

The *policy function* $h : \Re_+ \to \Re_+$ is defined by

$$h(x) = \hat{x}_1$$

where $\{\hat{x}_t\}_0^\infty$ is the optimal program from $x \in \Re_+$.

The properties of the value and policy functions can be summarized in the following result. This is based on Dutta and Mitra (1989) and Stokey, Lucas and Prescott (1989).

Proposition 3.1: *(i) The value function V is strictly concave and continuous on $\Re_+$ and non-decreasing on Y. Further, V is the unique continuous function on $Y \equiv [0, \xi]$ which satisfies the functional equation of dynamic programming:*

$$V(x) = \max_{y \in \Gamma(x)} [u(x, y) + \delta V(y)]$$

(ii) The policy function h satisfies the following property: for each $x \in \Re_+$, $h(x)$ is the unique solution to the constrained maximization problem:

$$\begin{array}{ll} \text{Maximize} & u(x, y) + \delta V(y) \\ \text{Subject to} & y \in \Gamma(x) \end{array}$$

Further, h is continuous on $\Re_+$.

In view of the definition of the policy function h, the optimal program from $x \in X$ is the trajectory $\{h^t(x)\}_0^\infty$ generated by the policy function. Thus, an optimal program from $x \in X$ can be called *periodic* (with period k) if x is a periodic point of h (with period k).

4. Duality Theory

4.1 Price Characterization of Optimality

Optimality can be characterized in terms of dual variables or shadow prices. The basic result of the theory, describing this characterization, can be stated as follows. [A full discussion can be found in Weitzman (1973) and McKenzie (1986)].

Proposition 4.1: (a) *If* $\{x_t\}_0^\infty$ *is an optimal program from* $\mathbf{x} \in X$ *and* $\mathbf{x} > 0$, *and there is some* $(\bar{x}, \bar{y}) \in \Omega$ *with* $\bar{y} > 0$, *then there is a sequence* $\{p_t\}_0^\infty$ *of non-negative prices such that for* $t \geq 0$,

(i) $\delta^t V(x_t) - p_t x_t \geq \delta^t V(x) - p_t x$ *for all* $x \geq 0$

(ii) $\delta^t u(x_t, x_{t+1}) + p_{t+1} x_{t+1} - p_t x_t \geq \delta^t u(x, y) + p_{t+1} y - p_t x$ *for all* $(x, y) \in \Omega$

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

(b) *If* $\{x_t\}_0^\infty$ *is a program from* $\mathbf{x} \geq 0$, *and there is a sequence* $\{p_t\}_0^\infty$ *of non-negative prices such that for* $t \geq 0$, (ii) *and* (iii) *above are satisfied, then* $\{x_t\}_0^\infty$ *is an optimal program from* $\mathbf{x}$.

If $\{x_t\}_0^\infty$ is a program from $\mathbf{x} \geq 0$, and $\{p_t\}_0^\infty$ is a non-negative sequence of prices satisfying (i), (ii) and (iii) of Proposition 4.1(a), we will say that the program $\{x_t\}_0^\infty$ is *price supported* by $\{p_t\}_0^\infty$. When $\{x_t\}_0^\infty$ is price supported by $\{p_t\}_0^\infty$, we refer to $\{p_t\}_0^\infty$ as a sequence of *present-value prices*. Associated with $\{p_t\}_0^\infty$ is a sequence $\{P_t\}_0^\infty$ of *current value prices* defined by

$$P_t = (p_t / \delta^t) \text{ for } t \geq 0$$

If the value function has finite steepness at zero, it is possible to choose the current value prices associated with any optimal program to be uniformly bounded above. To establish this, we proceed formally as follows. Normalize $u(0,0) = 0$, and note then that $V(0) = 0$. Then for all x in X, $(x, 0)$ is in Ω (by (A4)) and

$u(\zeta, 0) \geq u(x, 0) \geq 0$ by (A7). Further, for all x in X with $x > 0$, $V(x) \geq 0$ and $[V(x)/x]$ is decreasing in x by Proposition 3.1. We define

$$\rho \equiv \lim_{x \to 0+} [V(x)/x]$$

In the above definition, ρ can be infinite. If ρ happens to be finite, we say that V has "finite steepness".

Proposition 4.2: *Let (Ω, u, δ) be a dynamic optimization model. Suppose there is $(\hat{x}, \hat{y}) \in \Omega$ such that $\hat{x} \in X$ and $\hat{y} > 0$. Further, suppose that*

$$\rho \equiv \lim_{x \to 0+} [V(x)/x] < \infty \tag{4.1}$$

If $\{x_t\}_0^\infty$ is an optimal program from any $x \in X$, there is a price support $\{p_t\}_0^\infty$ of $\{x_t\}_0^\infty$ such that

$$P_t \equiv (p_t/\delta^t) \leq \beta \text{ for } t \geq 0 \tag{4.2}$$

where β is given by

$$\beta \equiv \max\ (\rho, [u(\zeta, 0) - u(\hat{x}, \hat{y}) + \rho\hat{x}]/\delta\hat{y}) \tag{4.3}$$

Proof: Consider, first, any optimal program $\{x_t\}_0^\infty$ from x in X, with $x > 0$. There are two cases to consider (i) $x_t > 0$ for all $t \geq 0$; (ii) $x_t = 0$ for some t.

In case (i), using Proposition 4.1 (i),

$$V(x_t) - P_t x_t \geq 0 \text{ for } t \geq 0$$

so that $P_t \leq [V(x_t)/x_t] \leq \rho \leq \beta$ for $t \geq 0$.

In case (ii), let T be the first period for which $x_t = 0$. Then $T \geq 1$, $x_t > 0$ for $t = 0, ..., T-1$, and $x_t = 0$ for $t \geq T$. By the argument used in case (i), $P_t \leq \rho$ for $t = 0, ..., T-1$. For $t = T$, using Proposition 4.1 (ii), we get

$$u(x_{T-1}, 0) - P_{T-1}x_{T-1} \geq u(\hat{x}, \hat{y}) + \delta P_T \hat{y} - P_{T-1}\hat{x}$$

so that $P_T \leq [u(\zeta, 0) - u(\hat{x}, \hat{y}) + \rho\hat{x}]/\delta\hat{y} \leq \beta$ by (4.3).

If $P_t \leq P_T$ for all $t > T$, then we are done. Otherwise, let τ be the first period ($> T$) for which $P_t > P_T$. Then $P_{\tau-1} \leq P_T < P_\tau$, and using Proposition 4.1 (ii), we get for all $(x, z) \in \Omega$,

$$0 \geq u(x, z) + \delta P_\tau z - P_{\tau-1}x \geq u(x, z) + \delta P_{\tau-1}z - P_{\tau-1}x.$$

Furthermore, since $x_{\tau-1} = 0$, we have

$$0 \geq V(x) - P_{\tau-1}x \text{ for all } x \geq 0$$

Thus, defining $P'_t = P_t$ for $t = 0, ..., \tau - 1$ and $P'_t = P_{\tau-1}$ for $t \geq \tau$, and $p'_t = \delta^t P'_t$ for $t \geq 0$, we see that $\{p'_t\}_0^\infty$ provides a price support to $\{x_t\}_0^\infty$, and $P'_t \leq \beta$ for $t \geq 0$.

The final step is to obtain a price support for the optimal program $\{0\}_0^\infty$ from 0, which satisfies (4.2). Since $\rho < \infty$, we have $V'_+(0) = \rho < \infty$. Thus, by concavity of V, we have

$$V(0) - P0 \geq V(x) - Px \text{ for all } x \geq 0 \tag{4.4}$$

where $P = V'_+(0) = \rho$. Then, by the induction argument of Weitzman (1973) [see also McKenzie (1986)], we can get a price support $\{p_t\}_0^\infty$ of the "zero program" $\{0\}_0^\infty$, with $P_0 = p_0 = P$, obtained in (4.4). Now, following the analysis of case (ii) above (identifying period T with period 0), we get a price support $\{p'_t\}_0^\infty$ of $\{0\}_0^\infty$, such that $P'_t \leq \beta$ for $t \geq 0$.

4.2 The Value-Loss Method

The value-loss method is based on the observation that at the prices supporting an optimal program, there is no activity which yields a higher "generalized profit" at any date (value of utility plus value of terminal stocks minus value of initial stocks at that date) than the activity chosen along the optimal program at that date. In other words, there are no arbitrage possibilities available at the supporting prices.

This observation leads to a basic tool for analyzing minimum impatience results (see Mitra (Chapter 11) for a proof) which we state in the following proposition.[3]

Proposition 4.3: *Let (Ω, u, δ) be a dynamic optimization model. Suppose $\{x_t\}_0^\infty$ is an optimal program with price support $\{p_t\}_0^\infty$, and $\{y_t\}_0^\infty$ is an optimal program with price support $\{q_t\}_0^\infty$. Denoting (p_t/δ^t) by P_t and (q_t/δ^t) by Q_t for $t \geq 0$, we have*

(i) $\delta(P_{t+1} - Q_{t+1})(y_{t+1} - x_{t+1}) \leq (P_t - Q_t)(y_t - x_t)$ *for* $t \geq 0$

(ii) $(P_t - Q_t)(y_t - x_t) \geq 0$ *for* $t \geq 0$

Furthermore, if $y_t \neq x_t$ for some t, then the inequalities in (i) and (ii) are strict for that t.

5. On a Relationship between Discounting and Complexity

5.1 The Montrucchio-Sorger Result

If we consider topological entropy to be an appropriate measure of "complexity" of a dynamical system, then a natural way to study the relationship between discounting and complicated optimal behavior is to find the relationship between the discount factor of a dynamic optimization model and the topological entropy of its policy function. This is the approach taken in Montrucchio (1994), where he establishes (under strong concavity assumptions on the utility function) that if (Ω, u, δ) is a dynamic optimization model with policy function h, A is a compact, invariant set contained in the interior of X, and $\psi(h, A)$ is the topological entropy of h on A, then the discount factor, δ, is related to the topological entropy by the inequality

$$\delta \leq e^{-\psi(h,A)} \tag{5.1}$$

Subsequently, it has been shown in Montrucchio and Sorger (1994) that the strong concavity assumption on the utility func-

[3]The inequalities of Proposition 4.3 must be very familiar objects from the point of view of the turnpike theory literature, where they suggest a natural choice of a Lyapunov function for the study of global asymptotic stability of optimal growth paths; see especially Brock and Scheinkman (1976), Cass and Shell (1976) and McKenzie (1986). The same inequalities have figured prominently in the literature on the intertemporal decentralization of the transversality condition; see especially Brock and Majumdar (1988) and Dasgupta and Mitra (1988).

tion can be dispensed with in deriving the inequality (5.1). In particular, it follows from their result that inequality (5.1) holds under the standard assumptions used in Section 3 of this chapter.

In this section, we will show how the relationship (5.1) (which I refer to henceforth as the Montrucchio-Sorger result) can be derived using the value-loss approach, thereby providing a unified view of the minimum impatience results obtained in the literature on chaotic optimal behavior.

In order to establish (5.1), we need two preliminary results. To describe the results, define $Z \equiv X - \{0\}$, and let A be a compact invariant set contained in Z. Given any x and y in A, let $\{p_t\}_0^\infty$ and $\{q_t\}_0^\infty$ be the price sequences supporting the optimal programs from x and y respectively. Denote by $\{P_t\}_0^\infty$ and $\{Q_t\}_0^\infty$ the "current" price sequences corresponding to $\{p_t\}_0^\infty$ and $\{q_t\}_0^\infty$ respectively, and to simplify notation denote P_0 by P and Q_0 by Q.

We know from Proposition 4.3 that

$$(P - Q)(y - x) \geq 0 \tag{5.2}$$

and the inequality in (5.2) is strict when $y \neq x$.

We need to establish, first, that there is $\beta > 0$, such that for all x, y in A,

$$(P - Q)(y - x) \leq \beta|x - y| \tag{5.3}$$

This amounts to establishing a uniform bound on the initial period supporting prices associated with optimal programs starting from initial stocks in A.

We also need to establish that given any $\varepsilon > 0$, there is $\alpha(\varepsilon) > 0$ such that

$$x, y \in A \text{ and } |x - y| \geq \varepsilon \text{ imply } (P - Q)(y - x) \geq \alpha(\varepsilon) \tag{5.4}$$

One recognizes this, of course, as a version of the well-known "value-loss lemma" appearing prominently in the turnpike literature since Radner (1961).

We now proceed to establish these two results formally.

Lemma 5.1: *Let A be a compact invariant set contained in Z. There is $\beta > 0$ such that for all x, y in A,*

$$(P - Q)(y - x) \leq \beta \; |x - y| \tag{5.5}$$

Proof: Given A, we can find $a \in Z$ such that $A \subset [a, \zeta]$. Denote $V'_{-}(a)$ by β.

Without loss of generality, suppose $y \geq x$. Then

$$(P - Q)(y - x) \leq P(y - x) \leq V'_{-}(x)(y - x) \tag{5.6}$$

Since $x \in A$, we have $V'_{-}(x) \leq V'_{-}(a) = \beta$. So (5.6) yields (5.5) immediately.

Lemma 5.2: *Let A be a compact invariant set contained in Z. For every $\varepsilon > 0$, there exists $\alpha(\varepsilon) > 0$ such that*

$$x, y \in A and \; |x - y| \geq \varepsilon imply \; (P - Q)(y - x) \geq \alpha(\varepsilon) \tag{5.7}$$

Proof: Given A, we can find $a \in Z$ such that $A \subset [a, \zeta]$. Denote $V'_{-}(a)$ by β.

If the Lemma were not true, there would exist a sequence (x^s, y^s) $s = 1, 2, \ldots$, with x^s, $y^s \in A$ and $|x^s - y^s|$ for all s, such that $(P^s - Q^s)(y^s - x^s) \to 0$ as $s \to \infty$.

Now, using (i) of Proposition 4.1, $0 \leq P^s \leq V'_{-}(x^s) \leq V'_{-}(a) = \beta$, $0 \leq Q^s \leq V'_{-}(y^s) \leq V'_{-}(a) = \beta$, so we can find a subsequence s' of s, such that $|P^{s'}| \to \hat{P}$, $|Q^{s'}| \to \hat{Q}$, $x^{s'} \to \bar{x}$ and $y^{s'} \to \bar{y}$ as $s' \to \infty$.

Denoting $(\bar{y} + \bar{x})/2$ by $\bar{z}$, and using (i) of Proposition 4.1 again, $V(x^s) - P^s x^s \geq V(\overline{z}) - P^s \bar{z}$ for all $s \geq 1$. Thus, by the continuity of V,

$$V(\bar{x}) - \hat{P}\bar{x} \geq V(\bar{z}) - \hat{P}\bar{z} \tag{5.8}$$

Since $|x^{s'} - y^{s'}| \geq \varepsilon$ for all s', we have $|\bar{x} - \bar{y}| \geq \varepsilon$. Thus, using the strict concavity of V, we get

$$V(\bar{z}) > \frac{1}{2}V(\bar{x}) + \frac{1}{2}V(\bar{y}) \tag{5.9}$$

Using (5.8) and (5.9), we obtain

$$V(\bar{x}) - \hat{P}\bar{x} > V(\bar{y}) - \hat{P}\bar{y} \tag{5.10}$$

Using y^s and Q^s in place of x^s and P^s respectively in the above steps, we can obtain similarly

$$V(\bar{y}) - \hat{Q}\bar{y} > V(\bar{x}) - \hat{Q}\bar{x} \tag{5.11}$$

Combining (5.10) and (5.11) yields

$$(\hat{P} - \hat{Q})(\bar{y} - \bar{x}) > 0 \tag{5.12}$$

But since $(P^{s'} - Q^{s'})(y^{s'} - x^{s'}) \to 0$ as $s' \to \infty$, and $P^{s'} \to \hat{P}$, $Q^{s'} \to \hat{Q}$, $y^{s'} \to \bar{y}$, $x^{s'} \to \bar{x}$ as $s' \to \infty$, we have

$$(\hat{P} - \hat{Q})(\bar{y} - \bar{x}) = 0 \tag{5.13}$$

which contradicts (5.12) and establishes the result.

We are now in a position to state and prove (a version of) the Montrucchio-Sorger result.

Theorem 5.1: *Let (Ω, u, δ) be a dynamic optimization model and let $h : X \to X$ be its policy function. Assume that A is a compact subset of X which is contained in Z, and which is invariant under h. Then,*

$$\delta \leq e^{-\psi(h,A)}$$

Proof: Let n be a positive integer, ε a positive real number, and B an (n, ε)-separated subset of A. For every x, y in B with $x \neq y$, there exists $t \in \{0, 1, ..., n-1\}$ such that $|h^t(x) - h^t(y)| > \varepsilon$. Using Proposition 4.3, we have

$$\delta^t(P_t - Q_t)(y_t - x_t) \leq (P - Q)(y - x) \tag{5.15}$$

where $x_t = h^t(x)$, $y^t = h^t(y)$, $P = P_0$ and $Q = Q_0$.

Using Lemma 5.1, we get

$$(P - Q)(y - x) \leq \beta|x - y| \tag{5.16}$$

Using Lemma 5.2 and $|x_t - y_t| > \varepsilon$, we get

$$(P_t - Q_t)(y_t - x_t) \geq \alpha(\varepsilon) \tag{5.17}$$

since $(x_t, x_{t+1}, ...)$ is an optimal program from x_t with current value supporting prices $(P_t, P_{t+1}, ...)$, and a similar remark applies to y_t and Q_t.

Combining (5.15), (5.16), (5.17), we get

$$\delta^t \alpha(\varepsilon) \leq \beta|x - y| \tag{5.18}$$

and since $t \leq n - 1$, and $0 < \delta < 1$,

$$\delta^n \alpha(\varepsilon)/\beta < |x - y| \tag{5.19}$$

Since $B \subset A \subset X$, where $X = [0, \zeta]$, and B is an (n, ε)-separated set, the number of elements of B must satisfy the inequality:

$$\text{Card B } < [\beta\zeta/\delta^n \alpha(\varepsilon)] + 1 < [2\beta\zeta/\delta^n \alpha(\varepsilon)] \tag{5.20}$$

Using (5.20), we obtain

$$\begin{aligned}
\psi_{\in}(h, A) &= \limsup_{n\to\infty} (1/n) \log[2\beta\zeta/\delta^n \alpha(\varepsilon)] \\
&= \limsup_{n\to\infty} [(1/n) \log[2\beta\zeta/\alpha(\varepsilon)] + (1/n) \log(1/\delta)^n] \\
&\leq \limsup_{n\to\infty} (1/n) \log(1/\delta)^n = \log(1/\delta)
\end{aligned}$$

This implies that $\psi(h, A) = \lim_{\in\to 0} \psi_{\in}(h, A) \leq \log(1/\delta)$. This yields

$$e^{\psi(h,A)} \leq (1/\delta)$$

which is equivalent to (5.14).

Remark 5.1: Notice that if A is a compact subset of X which is contained in the interior of X, and which is invariant under h, then, by Theorem 5.1, we have

$$\delta \leq e^{-\psi(h,A)}$$

which is the Montrucchio-Sorger result for our aggregative framework.

5.2 Discounting and Topological Entropy

A possible difficulty in applying Theorem 5.1 of the previous section is in identifying suitable compact subsets of Z which are invariant under h. Notice that while X itself is clearly a compact set which is invariant under h, the Montrucchio-Sorger result does not imply that

$$\delta \leq e^{-\psi(h,X)} \tag{5.21}$$

and it is clear from our method of proof (and also from the proofs of Montrucchio (1994) and Montrucchio and Sorger (1994)) that this is not an easy extension of Theorem 1.

If, for example, h is the well-known logistic map ($h(x) = 4x(1-x)$ for $x \in X \equiv [0,1]$), then h is t*opologically transitive*; that is, for any pair of open sets U_1, U_2 in X, there exists a positive integer k such that $h^k(U_1) \cap U_2$ is non-empty. For a proof of this fact, see Devaney (1989, p. 51). Then, by Lemma 37 of Block and Coppel (1992, p. 155), every proper closed subset of $[0,1]$, which is invariant under h, has empty interior. That is, these sets are "thin", and *would not* include, for instance, any open intervals. They *would* include finite sets, for instance those consisting of the points of a periodic cycle (of periodicity exceeding 1), but clearly the topological entropy of h on such sets is zero. Thus, compact subsets of Z which are invariant under h, might not be the sets that we would necessarily want to focus on.

This difficulty seems especially acute in the one-dimensional case, and less so in the multi-dimensional setting, for which the results of Montrucchio (1994) and Montrucchio-Sorger (1994) were developed. For instance, if $h : \Re^2 \to \Re^2$ is the Henon map ($h_1(x_1, x_2) = 1 + x_2 - 1.4\, x_1^2$, $h_2(x_1, x_2) = 0.3\, x_1$) then, as Henon (1976, p. 76) demonstrates, there is a rectangular region in the interior of $\Re^2$ which is mapped into itself by h.

In this section, we show under an additional assumption, how the formula (5.21) can be obtained. The extra assumption we use for this purpose involves "bounded steepness" of the reduced-form utility function, a concept introduced to the optimal growth literature in Gale (1967).

We now proceed formally as follows. Recall from Section 4 that we normalized $u(0,0) = 0$, so $V(0) = 0$. Then for all $x \in X$, $(x, 0)$ is in Ω (by (A4)) and $u(x, 0) \geq 0$ by (A7). Further, $[u(x,0)/x]$ is decreasing in x on X, by (A5). Our additional assumption is:

A8: $\sigma \equiv \lim_{x \to 0+} [u(x,0)/x] < \infty$

From this point onwards, we refer to a triplet (Ω, u, δ) as a dynamic optimization model if it satisfies assumptions (A1)-(A8).

We now proceed to establish the basic relationship between the discount factor and the topological entropy of the corresponding policy function.

Theorem 5.2: *Let (Ω, u, δ) be a dynamic optimization model and let $h : X \to X$ be its policy function. Then*

$$\delta \leq e^{-\psi(h,X)} \tag{5.22}$$

Proof: Notice, first, that if $h(x) = 0$ for all $x \in X$, then clearly $\psi(h, X) = 0$ and (5.22) holds trivially. Thus we concentrate on the situation where there is some $\hat{x} \in X$ such that $h(\hat{x}) > 0$.

Recall from Section 4 that we defined:

$$\rho \equiv \lim_{x \to 0+} [V(x)/x]$$

We subdivide our proof into two cases: (i) $\rho < \infty$ (ii) $\rho = \infty$.

Case I: $[\rho < \infty]$

In this case, defining β by (4.3), and applying Proposition 4.2, if $\{x_t\}_0^\infty$ is any optimal program from $x \in X$, then there is a price support $\{p_t\}_0^\infty$ of $\{x_t\}_0^\infty$ such that

$$P_t \equiv (p_t/\delta^t) \leq \beta \text{ for } t \geq 0$$

Then, we can follow the proofs of Lemma 5.1, Lemma 5.2 and Theorem 5.1 to obtain (5.22), replacing A by X in the appropriate steps of the proofs.

Case II: $[\rho = \infty]$

We can choose $0 < a < \zeta$, such that

$$[V(x)/x] > [\sigma/(1-\delta)] \text{ for } 0 < x \leq a \tag{5.23}$$

where σ is given by (A8).

We claim, first, that

$$h(x) > x \text{ for } 0 < x \leq a \tag{5.24}$$

For, suppose, on the contrary that there is some $0 < x \leq a$, for which $h(x) \leq x$. Then, using Proposition 3.1, $V(x) = u(x, h(x)) + \delta V(h(x)) \leq u(x, h(x)) + \delta V(x)$, so that

$$V(x) \leq u(x, h(x))/(1-\delta)$$

Using (A7), $u(x, h(x)) \leq u(x, 0)$, and we get

$$[V(x)/x] \leq \frac{[u(x,0)/x]}{(1-\delta)} \leq \frac{\sigma}{(1-\delta)}$$

which contradicts (5.23) and establishes (5.24).

Second, we claim that

$$h(x) > 0 \text{ for } 0 < x \leq \zeta \tag{5.25}$$

Suppose, on the contrary, there is some $0 < x \leq \zeta$ such that $h(x) = 0$. Then $\{x_t\}_0^\infty$ given by $(x, 0, 0, ...)$is the optimal program from x. Given (5.24), Proposition 4.1 can be used to get a price support $\{p_t\}_0^\infty$ of $\{x_t\}_0^\infty$. Denoting (p_t/δ^t) by P_t for $t \geq 0$, we have

$$V(x_1) - P_1 x_1 \geq V(y) - P_1 y \text{ for all } y \geq 0$$

Since $x_1 = 0$, and $V(0) = 0$, we get

$$V(y) \leq P_1 y \text{ for all } y \geq 0 \tag{5.26}$$

But, by letting $y \to 0$ in (5.26), we then contradict the fact that $\rho = \infty$. This establishes (5.25).

Since h is continuous on $[a, \zeta]$, (5.25) implies that there is $b' > 0$ such that

$$h(x) \geq b' \text{ for all } x \in [a, \zeta] \tag{5.27}$$

Define $b = min[a, b']$ and $A = [b, \zeta]$. Then, we claim that A is a compact, invariant set. Since compactness is clear, we proceed to check the invariance property. We have either (I) $b = a$, or (II) $b \neq a$. If $b = a$, then $b' \geq a = b$, and (5.27) implies that $h(x) \geq b$ for $x \in [b, \zeta]$. If $b \neq a$, then $b = b' < a$, and (5.27) implies that $h(x) \geq b$ for all $x \in [a, \zeta]$. For $x \in [b, a)$, we have $h(x) > x$ by (5.24), and so $h(x) \geq b$. Thus for all $x \in [b, \zeta]$, we get $h(x) \geq b$. Since $h(x) \leq \zeta$ for all $x \in [b, \zeta]$, we have established that $A = [b, \zeta]$ is an invariant set.

If $0 < x < b$, then there is some T large enough such that $h^T(x) \in A$ and so $h^t(x) \in A$ for $t \geq T$. If $x \in A$, then $h^t(x) \in A$ for $t \geq 0$. If $x = 0$, then $h^t(x) = 0$ for $t \geq 0$.

Define $C = \bigcap_{t=0}^{\infty} h^t(X)$. Then C is compact and invariant, and $C = C_1 \cup C_2$ where $C_1 = \bigcap_{t=0}^{\infty} h^t(A)$ and $C_2 = \{0\}$. Clearly, C_1 and C_2 are compact, invariant sets.

Now, $\psi(h, X) = \psi(h, C)$ by Corollary 4.1.8, p. 196 of Alseda, Llibre and Misiurewicz (1993). Further, $\psi(h, C_2) = 0$, since from the definition of topological entropy, if W is any finite set, then the topological entropy of any map $g : W \to W$ is zero. Thus, using Lemma 4.1.10, p. 197 of Alseda, Llibre and Misiurewicz (1993), $\psi(h, C) = \psi(h, C_1)$. Summarizing, we have $\psi(h, X) = \psi(h, C_1)$.

Since $C_1 \subset A$, we note that C_1 is a compact, invariant set in Z, and so by Theorem 5.1,

$$\delta \leq e^{-\psi(h, C_1)}$$

Since $\psi(h, C_1) = \psi(h, X)$, we have established (5.22).

Remark 5.2: We have, so far, not been able to construct a proof of Theorem 5.2, which dispenses with assumption (A8). However, this assumption establishes the strong claim in (5.24), and that does not seem to be necessary for the rest of the proof in case (ii) to be valid.

5.3 Discounting and Metric Entropy

We have already indicated in Section 2.6 that topological chaos may not be observable, and so topological entropy might not be an appropriate measure of the complexity of a dynamical system. In this context, a natural alternative measure to consider is the metric (or measure-theoretic) entropy of a dynamical system, and to conclude that the system exhibits complicated behavior when the metric entropy is positive.

Interestingly, if we adopt this point of view, our analysis so far is still seen to be extremely useful in studying the relation between discounting and complicated behavior of a dynamical system. This is clear by noting the basic relationship between topological and metric entropy, due to Goodwyn (1968).[4]

Proposition 5.1: *Let (f, I) be a dynamical system (as in Section 2.1), and $M(f, I)$ be the set of all f-invariant probability measures on the Borel sets of I. If $\mu \in M(f, I)$, then*

$$\Phi_\mu(f, I) \leq \psi(f, I) \tag{5.28}$$

In view of Proposition 5.1 and Theorem 5.2, we have the following result relating discounting and metric entropy.

Theorem 5.3: *Let (Ω, u, δ) be a dynamic optimization model and let $h : X \to X$ be its policy function. If μ is an h-invariant probability measure on the Borel sets of X, then*

$$\delta \leq e^{-\Phi_\mu(h,X)} \tag{5.29}$$

---

[4]The result of Goodwyn (1968) is applicable to more general dynamical systems than ours. It turns out that the topological entropy, $\psi(f, I)$, is the *supremum* over all $\mu \in M(f, I)$ of the metric entropies, $\Phi_\mu(f, I)$. This was first established by Dinaburg (1970). For a discussion of these results in the most general setting, see Goodman (1971).

Remark 5.3: Nishimura, Sorger and Yano (Chapter 9) have provided an example in which for every $0 < \delta < 1$, there is a dynamic optimization model (Ω, u, δ) such that the policy function, h_δ, exhibits ergodic chaos. However, as they have noted, the metric entropy of h_δ converges to zero as δ converges to 1. Theorem 5.3 shows that this is true not only for that example but in general; that is, if $(\Omega_s, u_s, \delta_s)$ is a sequence of dynamic optimization models with policy functions $h_s(s = 1, 2, ...)$, and $\delta_s \to 1$ as $s \to \infty$, then the metric entropy of h_s must converge to zero at $s \to \infty$.

6. Discount Factor Restrictions for Turbulence and Topological Chaos

In this section, we demonstrate the usefulness of the Montrucchio-Sorger result, by deriving a number of implications of it. Specifically, we show how the result can be used to derive discount factor restrictions for policies exhibiting turbulence and topological chaos.

The basic technical background that we need to study discount factor restrictions for periodic programs is a result due to Block, Guckenheimer, Misiurewicz and Young (1980) which provides a formula for the topological entropy of any continuous function exhibiting a periodic program with period not equal to a power of 2. We state this result here for ready reference.

Proposition 6.1: *Let $f : I \to I$ be a continuous map with an orbit of period $q = np$ where $n > 1$ is odd and $p = 2^k$ with $k \geq 0$. Then*

$$\psi(f) \geq (\log\ \lambda_n)/p \tag{6.1}$$

where λ_n is the unique positive root of the equation

$$z^n - 2z^{n-2} - 1 = 0 \tag{6.2}$$

If we combine Proposition 6.1 with Theorem 5.2, we obtain the following result immediately.

Theorem 6.1: *Let (Ω, u, δ) be a dynamic optimization model which exhibits a periodic program of period $q = np$ where $n > 1$ is* *odd and* $p = 2^k$ *with* $k \geq 0$. *Then*

$$\delta \leq 1/\lambda_n^{1/p} \tag{6.3}$$

where λ_n *is the unique positive root of the equation*

$$z^n - 2z^{n-2} - 1 = 0 \tag{6.4}$$

Given Theorem 6.1, all one needs to obtain suitable discount factor restrictions for policy functions which exhibit periodic programs of positive topological entropy (and, therefore, which exhibit topological chaos) is an accurate calculation of the (unique) positive root of the polynomial in (6.4).

We illustrate this point with the simplest case, where the dynamic optimization model (Ω, u, δ) exhibits a period-three cycle. Here, of course, $n = 3$, $p = 1$ (so $k = 0$) and $q = np = 3$.

The polynomial in (6.4) reduces to

$$z^3 - 2z - 1 = 0 \tag{6.5}$$

It is easy to verify that

$$\lambda_3 = [\sqrt{5} + 1]/2 \tag{6.6}$$

is the unique positive root of (6.5). Thus, the discount factor restriction for a period-three cycle, by applying Theorem 6.1, is

$$\delta \leq 1/\lambda_3 \tag{6.7}$$

Now, the magnitude $(1/\lambda_3)$ can be written as

$$\begin{aligned} 2/(\sqrt{5}+1) &= 2(\sqrt{5}-1)/(\sqrt{5}+1)(\sqrt{5}-1) \\ &= 2(\sqrt{5}-1)/(5-1) \\ &= (\sqrt{5}-1)/2 \end{aligned}$$

This calculation leads to the following Corollary, which was first established by Sorger (1992b) by using entirely different methods.

Corollary 6.1: *Let* (Ω, u, δ) *be a dynamic optimization model, which exhibits a period-three cycle. Then*

$$\delta \leq (\sqrt{5} - 1)/2 \tag{6.8}$$

In order to obtain a discount factor restriction for a turbulent policy function, we need the following mathematical result.

Proposition 6.2: *Let* $f : I \to I$ *be a continuous map. Let* $J_1, \ldots, J_p$ *be closed intervals with pairwise disjoint interiors and let* $A = (a_{ik})$ *be the* $p \times p$ *matrix defined by*

$$a_{ik} = \begin{cases} 1 & \text{if } J_k \subset f(J_i) \\ 0 & \text{otherwise} \end{cases}$$

Then $\psi(f) \geq \log \lambda$, *where* λ *is the maximal eigenvalue of* A.

If we combine Proposition 6.2 with Theorem 5.2, we obtain the following result immediately.

Theorem 6.2: *Let* (Ω, u, δ) *be a dynamic optimization model, with policy function,* $h : X \to X$. *If* h *exhibits turbulence, then*

$$\delta \leq (1/2) \tag{6.9}$$

Proof: If h exhibits turbulence, then we can find closed subintervals J_1 and J_2 of X, with at most one point in common, such that

$$J_1 \cup J_2 \subset f(J_1) \cap f(J_2)$$

Then the interiors of J_1 and J_2 are disjoint, and the matrix $A = (a_{ik})$ of Proposition 6.1 is

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

The maximal eigenvalue of A is 2. So, applying Proposition 6.1, we get

$$\psi(f) \geq \log\ 2 \tag{6.10}$$

Using (6.10) and Theorem 5.2, we get (6.9).

We can use Theorem 6.1 to obtain an upper bound on the discount factor, δ, that must be satisfied in order that a dynamic optimization model (Ω, u, δ) yield a periodic optimal program of odd period greater than one.

Corollary 6.2: *Suppose (Ω, u, δ) is a dynamic optimization model with policy function, h. Let $n > 1$ be any odd integer. If h has a periodic orbit of period n, then*

$$\delta < (1/\sqrt{2}) \tag{6.11}$$

Proof: Using Theorem 6.1, we know that

$$\delta \leq (1/\lambda_n) \tag{6.12}$$

where λ_n is the unique positive root of the equation

$$z^n - 2z^{n-2} - 1 = 0 \tag{6.13}$$

Define $g(z) = z^n - 2z^{n-2} - 1$ for $z \geq 0$. For $z = \sqrt{2}$, $\{g(z)/z^{n-2}\} = z^2 - 2 - (1/z^{n-2}) < z^2 - 2 = 0$. And if $z \geq \sqrt{2 + \{1/(\sqrt{2})^{n-2}\}}$, then $z^2 - 2 - (1/z^{n-2}) > z^2 - 2 - \{1/(\sqrt{2})^{n-2}\} \geq 2 + \{1/(\sqrt{2})^{n-2}\} - 2 - \{1/(\sqrt{2})^{n-2}\} = 0$. Thus we know that

$$\sqrt{2} < \lambda_n < \sqrt{2 + \{1(\sqrt{2})^{n-2}\}} \tag{6.14}$$

Combining (6.12) and (6.14), we get (6.11).

More generally, Proposition 6.1, Theorem 5.2, and (6.14) can be used to obtain upper bounds on discount factors that must hold in order that optimal programs be periodic with period $q = np$ where $n > 1$ is odd and $p = 2^k$ with $k \geq 0$.

Corollary 6.3: *Suppose (Ω, u, δ) is a dynamic optimization model with policy function, h. Let $n > 1$ be an odd integer, k be a non-negative integer and $q = n2^k$. If h has a periodic orbit with period q, then*

$$\delta < (1/\sqrt{2})^{(1/2^k)} \tag{6.15}$$

7. An Open Question

There are two results which provide "exact" discount factor restrictions for chaotic behavior in aggregate dynamic optimization models, the first dealing with the case of period-three cycles (see Mitra (Chapter 11) and Nishimura and Yano (Chapter 12)), and the second dealing with turbulence (see Mitra (1994b)). We state them here in order to provide a self-contained discussion.[5]

Proposition 7.1: *Let (Ω, u, δ) be a dynamic optimization model, with a policy function h. If h exhibits a period-three cycle, then*

$$\delta < [(\sqrt{5} - 1)/2]^2 \tag{7.1}$$

Conversely, if δ satisfies (7.1), then one can construct a transition possibility set, Ω, and a reduced form utility function u, such that the dynamic optimization model (Ω, u, δ) has an optimal program exhibiting a period-three cycle.

Proposition 7.2: *Let (Ω, u, δ) be a dynamic optimization model, with a policy function h. If h exhibits turbulence, then*

$$\delta < (1/4) \tag{7.2}$$

Conversely, if δ satisfies (7.2), then one can construct a transition possibility set, Ω, and a reduced form utility function, u, such that the dynamic optimization model (Ω, u, δ) has a policy function which exhibits turbulence.

The observation we would like to make is that neither the discount factor restriction (7.1) for period-three cycles nor the discount factor restriction (7.2) for turbulence can be obtained from Theorem 5.2 which provides the basic relationship between the discount factor and the topological entropy of the corresponding policy function. In fact, as we have seen in Section 6, Theorem 5.2

[5] These results were proved for dynamic optimization models satisfying (A.1)-(A.7). Clearly, the "necessity part" of the results remain valid for dynamic optimization models satisfying (A.1)-(A.8). Furthermore, it can be checked that the examples used to demonstrate the "sufficiency part" satisfy (A.8), besides (A.1)-(A.7).

yields a discount factor restriction of $(\sqrt{5}-1)/2$ for period-three cycles (Theorem 6.1) and a discount factor restriction of $(1/2)$ for turbulence (Theorem 6.2). Interestingly, these restrictions are exactly the square roots of the exact restrictions appearing in (7.1) and (7.2) respectively.

The open question which this observation naturally raises is whether there is a more refined relationship (than is provided in Theorem 5.2) between the discount factor and the topological entropy, which would yield the exact discount factor restrictions for period-three cycles and turbulence [in (7.1) and (7.2) respectively] as special cases.