ELSEVIER

Journal of Economic Behavior & Organization
Vol. 33 (1998) 421–434

JOURNAL OF
Economic Behavior
& Organization

On the relationship between discounting and complicated behavior in dynamic optimization models

Tapan Mitra*

Department of Economics, 418 Uris Hall, Cornell University, Ithaca NY 14853, USA

Received 18 January 1996; accepted 17 March 1997

Abstract

This paper explores the relationship between discounting and complicated behavior in dynamic optimization models. An inverse relationship between the discount factor of a dynamic optimization model and the topological entropy of the corresponding optimal policy function is established. For an aggregative model, this generalizes a result obtained by Montrucchio and Sorger. The generalization makes it possible to apply directly the results developed by Block, Guckenheimer, Misiurewicz and Young on the computation of the topological entropy of dynamical systems to obtain upper bounds on the discount factor necessary for the occurrence of topological chaos in aggregative dynamic optimization models. © 1998 Elsevier Science B.V.

JEL classification: C61; E32; O41

Keywords: Optimization; Dynamics; Discounting; Entropy

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 the complicated behavior generated by the corresponding (optimal) policy function has been a topic of extensive study.

* Corresponding author.

PII S0167-2681(97)00067-X

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 for exhibiting complicated behavior, the dynamic optimization model 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 factor, δ, must satisfy: $\delta < (\sqrt{5}-1)/2 \approx 0.618$. In subsequent work, Sorger (1994) refined the above bound to $\delta<0.5479$. The bound was further refined to

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

in Mitra (1996), Nishimura and Yano (1996). Furthermore, the bound in Eq. (1) 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. 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 (1996)) that, combining Eq. (1) with the "Sarkovskii order"[1], the following bound on the discount factor can be obtained:

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

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

$$\delta \leq 1/e^{\psi(h,A)} \tag{3}$$

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 (1996) have shown that the inequality holds even when the strong concavity assumption on the utility

[1] For this, and other basic concepts and results on chaotic dynamical systems, see, for example, Block and Coppel (1992), Day (1994), Devaney (1989).

function is replaced by a relatively mild strict concavity assumption on the utility function.

In this paper, we do three things. First, we establish the Montrucchio–Sorger result (Eq. (3)) 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 use for the study of complicated optimal behavior as well, and the inequality (Eq. (3)) is seen to rest firmly on a version of the well-known 'value-loss lemma,' which was introduced to turnpike theory by Radner (1961), and which has occupied center-stage in this literature thereafter.

Second, under an additional assumption of 'bounded steepness' of the utility function (see Section 4 for details) we extend the Montrucchio–Sorger result to establish that

$$\delta \leq 1/e^{\psi(h,X)} \tag{4}$$

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 Eq. (4).

Third, I explore the implications of result Eq. (4) for periodic 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 behavior, by Block et al. (1980). We use these results to show (among other things) that the inequality (Eq. (4)) yields the discount factor restriction: $\delta \leq (\sqrt{5}-1)/2$ for period three cycles.

2. Chaos

Let I be a compact interval in $\Re$, the set of reals. Let f:$I\rightarrow I$ 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 $\gamma(x)=\{y\text{: } y=f^n(x) \text{ for some } n\geq 0\}$.

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.

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

$$\psi_\varepsilon(f,A)=\lim_{n\rightarrow\infty}\sup(1/n)\log s(n,\varepsilon)$$

and the *topological entropy*[2] of f as

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

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).

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 $\Gamma(x)=\{y\in R_+: (x,y) \in \Omega\}$ for each $x \in \Re_+$.

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

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

If one is in state x today and moves to state z tomorrow (with $(x,z) \in \Omega$) then there is an immediate utility 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 following assumptions are imposed on Ω and u:

(A.1) (i) $(0,0) \in \Omega$, (ii) $(0,z) \in \Omega$ implies $z=0$.

(A.2) Ω is (i) closed, and (ii) convex.

(A.3) There is $\xi>0$ such that $(x,z) \in \Omega$ and $x \geq \xi$ implies $z < x$.

(A.4) 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 $x\leq\zeta$, 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:

(A.5) u is concave on Ω; further if (x,z) and (x',z') are in Ω, and $x \neq x'$, then for every

$$0 < \lambda < 1,\ u(\lambda(x,z) + (1-\lambda)(x',z')) > \lambda u(x,z) + (1-\lambda)u(x',z').$$

(A.6) u is upper semi-continuous on Ω.

(A.7) If $x, x'\in Y$, $(x,z)\in\Omega$, $x'\geq x$ and $0\leq z'\leq z$, then $u(x',z')\geq u(x,z)$.

[2] The formal definition of topological entropy was given by Adler et al. (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).

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 (A.1)–(A.7), 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 et al. (1989).

Proposition 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:* $\text{Max}_{y \in \Gamma(x)}[u(x,y) + \delta V(y)]$. *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).

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

1. $\delta^t V(x_t) - p_t x_t \geq \delta^t V(x) - p_t x$ *for all* $x \geq 0$, $t \geq 0$
2. $\delta^t u(x_t, x_{t+1}) + p_{t+1} x_{t+1} - p_t x_t \geq \delta^t u(x,y) + p_{t+1} y - p_t x$ *for all* $(x,y) \in \Omega$, $t \geq 0$
3. $\lim_{t \to \infty} p_t x_t = 0$

(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,* (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 7(a), we will say that the program $\{x_t\}_0^\infty$ is *price supported* by $\{p_t\}_0^\infty$. In this case, we refer to $\{p_t\}_0^\infty$ as a sequence of *present-value prices*, and associate with it a sequence $\{P_t\}_0^\infty$ of *current-value prices* defined by $P_t=(p_t/\delta^t)$ 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, a result which we state in Proposition 3 below (the proof is in the mathematical Appendix A). To this end, normalize $u(0,0)=0$, and note that $V(0)=0$. Then for all x in X, $(x,0)$ is in Ω (by (A.4)) and $u(\zeta,0)\geq u(x,0)\geq 0$ by (A.7). Further, for all x in X with $x>0$, $V(x)\geq 0$ and $[V(x)/x]$ decreases in x by Proposition 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 3. *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 $\rho\equiv\lim_{x\to 0^+}[V(x)/x]<\infty$. 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 $P_t\equiv(p_t/\delta^t)\leq\beta$* for $t\geq 0$, *where β is given by* $\beta\equiv\max(\rho,[u(\zeta,0)-u(\hat{x},\hat{y})+\rho\hat{x}]/\delta\hat{y})$.

3.4. 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 (1996) for a proof) which we state in the following proposition.

Proposition 4. *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*

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

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

4. On a relationship between discounting and complexity

4.1. 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

$$\delta \leq e^{-\psi(h,A)} \tag{5}$$

Subsequently, it has been shown in Montrucchio and Sorger (1996) that the strong concavity assumption on the utility function can be dispensed with in deriving the inequality (Eq. (5)). In particular, it follows from their result that inequality (Eq. (5)) holds under the standard assumptions used in Section 3 of this paper.

In this section, we will show how the relationship (Eq. (5)) (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 (Eq. (5)), 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 that $(P-Q)(y-x) \geq 0$ and this inequality is strict when $y \neq x$. We need to establish, first, that there is $\mu > 0$, such that for all $x,y \in A$, $(P-Q)(y-x) \leq \mu|x-y|$. 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 $x,y \in A$ and $|x-y| \geq \varepsilon$ imply $(P-Q)(y-x) \geq \alpha(\varepsilon)$. 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 state these two results formally; proofs are provided in the mathematical Appendix A.

Lemma 1. *Let A be a compact invariant set contained in Z. There is $\mu > 0$ such that for all x,y in A, $(P-Q)(y-x) \leq \mu|x-y|$.*

Lemma 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 $(P-Q)(y-x) \geq \alpha(\varepsilon)$.

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

Theorem 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 $t \in \{0,1,\ldots,n-1\}$ such that $|h^t(x)-h^t(y)| > \varepsilon$. Using Proposition 4, we have $\delta^t(P_t - Q_t)(y_t - x_t) \leq (P-Q)(y-x)$ where $x_t = h^t(x)$, $y^t = h^t(y)$, $P = P_0$ and $Q = Q_0$. Using Lemma 1, we get $(P-Q)(y-x) \leq \mu|x-y|$. Using Lemma 2 and $|x_t - y_t| > \varepsilon$, we get $(P_t - Q_t)(y_t - x_t) \geq \alpha(\varepsilon)$ since

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

Combining the above three inequalities, we get $\delta^t\,\alpha(\varepsilon) \le \mu|x-y|$ and since $t \le n-1$, and $0<\delta<1$, $\delta^n\,\alpha(\varepsilon)/\mu < |x-y|$. 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: Card B $< [\mu\zeta/\delta^n\alpha(\varepsilon)] + 1 < [2\mu\zeta/\delta^n\alpha(\varepsilon)]$. This yields

$$\psi_\in(h,A) = \limsup_{n\to\infty}(1/n)\log[2\mu\zeta/\delta^n\alpha(\varepsilon)] = \limsup_{n\to\infty}[(1/n)\log[2\mu\zeta/\alpha(\varepsilon)] + (1/n)\log(1/\delta)^n]$$
$$\le \limsup_{n\to\infty}(1/n)\log(1/\delta)^n = \log(1/\delta)$$

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

Remark. Note 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 1, we have $\delta \le e^{-\psi(h,A)}$ which is the Montrucchio–Sorger result for our aggregative framework.

4.2. *Discounting and topological entropy*

A possible difficulty in applying Theorem 1 of the previous section is in identifying suitable compact subsets of Z which are invariant under h. Note that while X itself is clearly a compact set which is invariant under h, the Montrucchio–Sorger result does not imply that

$$\delta \le e^{-\psi(h,X)} \tag{6}$$

and it is clear from our method of proof (and also from the proofs of Montrucchio (1994) and Montrucchio and Sorger (1996)) 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 *topologically 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.

In this section, we show under an additional assumption, how the formula (Eq. (6)) can be obtained. The extra assumption we use for this purpose involves 'bounded steepness' of the utility function, a concept introduced to the optimal growth literature in Gale (1967).

We now proceed formally as follows. Recall from Section 3 that we normalized $u(0,0) = 0$, so $V(0) = 0$. Then for all $x \in X$, $(x,0)$ is in Ω (by (A.4)) and $u(x,0) \ge 0$ by (A.7). Further, $[u(x,0)/x]$ decreases in x on X, by (A.5). Our additional assumption is:

(A.8) $\sigma \equiv \lim_{x\to 0^+}[u(x,0)/x] < \infty$

From this point onward, we refer to a triplet (Ω,u,δ) as a *dynamic optimization model* if it satisfies assumptions (A.1)–(A.8).

Theorem 2. *Let (Ω,u,δ) be a dynamic optimization model and let $h: X \to X$ be its policy function. Then*

$$\delta \leq e^{-\psi(h,X)} \tag{7}$$

Proof. Note, first, that if $h(x)=0$ for all $x \in X$, then clearly $\psi(h,X)=0$ and (Eq. (7)) 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 3 that we defined: $\rho \equiv \lim_{x \to 0^+}[V(x)/x]$. We consider two cases: (i) $\rho<\infty$ (ii) $\rho=\infty$.

Case (i) $[\rho<\infty]$ In this case, defining β as in Proposition 3, 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 $P_t \equiv (p_t/\delta^t) \leq \beta$ for $t \geq 0$. Then, we can follow the proofs of Lemma 1, Lemma 2 and Theorem 1 to obtain (Eq. (7)), replacing A with X in the appropriate steps of the proofs.

Case (ii) $[\rho = \infty]$ We can choose $0<a<\zeta$, such that $[V(x)/x]>[\sigma/(1-\delta)]$ for $0<x\leq a$ where σ is given by (A.8).

We claim, first, that $h(x)>x$ for $0<x\leq a$. Suppose, on the contrary that there is some $0<x\leq a$, for which $h(x)\leq x$. Then, using Proposition 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 (A.7), $u(x,h(x))\leq u(x,0)$, and we get $[V(x)/x]\leq[u(x,0)/x]/(1-\delta)\leq\sigma/(1-\delta)$ a contradiction which establishes the first claim.

Second, we claim that $h(x)>0$ for $0<x\leq\zeta$. Suppose, on the contrary, there is some $0<x \leq \zeta$ such that $h(x)=0$. Then $\{x_t\}_0^\infty$ given by $(x,0,0,\ldots)$ is the optimal program from x. Then Proposition 2 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 $V(x_1)-P_1 x_1 \geq V(y)-P_1 y$ for all $y\geq 0$. Since $x_1=0$, and $V(0)=0$, we get $V(y)\leq P_1 y$ for all $y\geq 0$. But, by letting $y\to 0$, we then contradict the fact that $\rho=\infty$. This establishes the second claim.

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

$$h(x) \geq b' \text{ for all } x \in [a,\zeta] \tag{8}$$

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 (Eq. (8)) implies that $h(x)\geq b$ for $x\in[b,\zeta]$. If $b\neq a$, then $b=b'<a$, and (Eq. (8)) implies that $h(x)\geq b$ for all $x\in[a,\zeta]$. For $x\in[b,a)$, we have $h(x)>x$, 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 $C=\cap_{t=0}^\infty h^t(X)$. Then C is compact and invariant, and $C=C_1\cup C_2$ where $C_1=\cap_{t=0}^\infty h^t(A)$ 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, (Alseda et al. (1993) (p.196)). 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, (Alseda et al. (1993) (p.197)), $\psi(h,C) =\psi(h,C_1)$. 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 1, $\delta \leq e^{-\psi(h,C_1)}$. Since $\psi(h,C_1)=\psi(h,X)$, we have established (Eq. (8)).

4.3. Discounting and metric entropy

Topological chaos may not be observable, 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.[3] Let (f,I) be a dynamical system (as in Section 2), and $M(f,I)$ be the set of all f-invariant probability measures on the Borel sets of I. If $\nu \in M(f,I)$, then $\Phi_\nu(f,I) \leq \psi(f,I)$.

Theorem 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*

$$\delta \leq e^{-\Phi_\nu(h,X)} \tag{9}$$

Remark. Nishimura et al. (1994) 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 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 $h_s (s=1,2,\ldots)$, and $\delta_s \to 1$ as $s \to \infty$, then the metric entropy of h_s must converge to zero at $s \to \infty$.

5. Discount factor restrictions for 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 topological chaos.

---

[3] For the definition of metric entropy, as well as this result, see Goodwyn (1969). The result of Goodwyn is applicable to more general dynamic systems than ours. It turns out that the topological entropy, $\psi(f,I)$, is the *supremum* over all $\nu \in M(f,I)$ of the metric entropies $\Phi_\nu(f,I)$. This was first established by Dinaburg (1970). For a discussion of these results in the most general setting, see Goodman (1971).

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

Proposition 5. *Let f:I→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 $\psi(f)\geq(\log \lambda_n)/p$ where λ_n is the unique positive root of the equation $z^n-2z^{n-2}-1=0$.*

Combining Proposition 5 with Theorem 2, we obtain the following result.

Theorem 4. *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 $\delta \leq 1/\lambda_n^{1/p}$ where λ_n is the unique positive root of the equation $z^n-2z^{n-2}-1=0$.*

Given Theorem 4, 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 equation: $z^n-2z^{n-1}-1=0$.

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 relevant polynomial is: $z^3-2z-1=0$. It is easy to verify that $\lambda_3=[\sqrt{5}+1]/2$ is the unique positive root of this polynomial. Thus, the discount factor restriction for a period three cycle, by applying Theorem 4, is $\delta\leq 1/\lambda_3$. Now, the magnitude $(1/\lambda_3)$ can be written as $(\sqrt{5}-1)/2$, and this leads to the following Corollary, which was first established by Sorger (1992b) by using entirely different methods.

Corollary 1. *Let (Ω,u,δ) be a dynamic optimization model, which exhibits a period three cycle. Then $\delta \leq (\sqrt{5}-1)/2$.*

We can also use Theorem 4 to obtain an upper bound on the discount factor, δ, that must be satisfied in order that a dynamic optimization model (Ω,u,δ) yields a periodic optimal program of odd period greater than one.

Corollary 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 $\delta < (1/\sqrt{2})$.*

Proof. 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

$$\sqrt{2}<\lambda_n<\sqrt{2+1/(\sqrt{2})^{n-2}\}} \tag{10}$$

Using Theorem 4, with $k=0$, and (Eq. (10)), we get $\delta<(1/\sqrt{2})$.

More generally, Proposition 5, Theorem 2, and (Eq. (10)) 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 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 $\delta < (1/\sqrt{2})^{(1/2^k)}$.*

Acknowledgements

I would like to thank Jess Benhabib, Kazuo Nishimura, Gerhard Sorger and Luigi Montrucchio, whose contributions on the relationship between discounting and optimal behavior have greatly influenced my own research on the topic. The present version has benefited from the discussion of the paper at the Viennese Workshop by Gerhard Sorger, and from comments by Richard Day and Lionel McKenzie.

Appendix A

Mathematical appendix

Proof of Proposition 3. Consider, first, any optimal program $\{x_t\}_0^\infty$ from x in X, with $x>0$. We consider two cases: (i) $x_t>0$ for all $t\geq 0$; (ii) $x_t=0$ for some t.

In case (i), using Proposition 2(i), $V(x_t)-P_t x_t \geq 0$ 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 $t=0,\ldots,T-1$, and $x_t=0$ for $t\geq T$. By the argument used in case (i), $P_t\leq\rho$ for $t=0,\ldots,T-1$. For $t=T$, using Proposition 2(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 $P_T\leq [u(\zeta,0)-u(\hat{x},\hat{y})+\rho\hat{x}]/\delta\hat{y}\leq\beta$.

If $P_t\leq P_T$ for all $t>T$, then we are done. Otherwise, let $\tau>T$ be the first period for which $P_t>P_T$. Then $P_{\tau-1}\leq P_T<P_\tau$, and using Proposition 2(ii), we get for all $(x,z)\in\Omega$, $0\geq u(x,z)+\delta P_\tau z-P_{\tau-1}x\geq u(x,z)+\delta P_{\tau-1}z-P_{\tau-1}x$. Also, since $x_{\tau-1}=0$, we have $0\geq V(x)-P_{\tau-1}x$ for all $x\geq 0$. Thus, defining $P'_t=P_t$ for $t=0,\ldots,\tau-1$ 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 $t\geq 0$.

It remains to obtain a price support for the optimal program $\{0\}_0^\infty$ from 0, with bounded current value prices. Since $\rho<\infty$, we have $V'_+(0)=\rho<\infty$. Thus, by concavity of V, we have $V(0)-P0\geq V(x)-Px$ for all $x\geq 0$, where $P=V'_+(0)$. 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 $P_0=p_0=P$. Now, following the analysis of case (ii) (identifying period T with period 1), we obtain a price support $\{p'_t\}_0^\infty$ of $\{0\}_0^\infty$, such that $P'_t\leq\beta$ for $t\geq 0$.

Proof of Lemma 1. Given A, we can find $a \in Z$ such that $A \subset [a,\zeta]$. Denote $V'_{-}(a)$ by μ. Without loss of generality, suppose $y \geq x$. Then $(P-Q)(y-x) \leq P(y-x) \leq V'_{-}(x)(y-x) \leq V'_{-}(a)(y-x) = \mu(y-x)$.

Proof of Lemma 2. Given A, we can find $a \in Z$ such that $A \subset [a,\zeta]$. Denote $V'_{-}(a)$ by μ. 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 $|x^s - y^s| \geq \varepsilon$ for all s, such that $(P^s - Q^s)(y^s - x^s) \to 0$ as $s \to \infty$.

Using Proposition 2(i), $0 \leq P^s \leq V'_{-}(x^s) \leq V'_{-}(a) = \mu$, and similarly, $0 \leq Q^s \leq \mu$, so we can find a subsequence s' of s, such that $P^{s'} \to \hat{P}, Q^{s'} \to \hat{Q}, x^{s'} \to \bar{x}, y^{s'} \to \bar{y}$ as $s' \to \infty$. Denoting $(\bar{y}+\bar{x})/2$ by $\bar{z}$, and using Proposition 2(i) again, $V(x^s) - P^s x^s \geq V(\bar{z}) - P^s \bar{z}$ for all $s \geq 1$. By continuity of V, $V(\bar{x}) - \hat{P}\bar{x} \geq V(\bar{z}) - \hat{P}\bar{z}$. Similarly, $V(\bar{y}) - \hat{Q}\bar{y} \geq V(\bar{z}) - \hat{Q}\bar{z}$.

Since $|x^{s'} - y^{s'}| \geq \varepsilon$ for all s', we have $|\bar{x} - \bar{y}| \geq \varepsilon$. Thus, using strict concavity of V, we get $V(\bar{z}) > (1/2)V(\bar{x}) + (1/2)V(\bar{y})$. Thus $V(\bar{x}) - \hat{P}\bar{x} > (1/2)V(\bar{x}) + (1/2)V(\bar{y}) - (1/2)\hat{P}\bar{x} - (1/2)\hat{P}\bar{y}$, so that $V(\bar{x}) - \hat{P}\bar{x} > V(\bar{y}) - \hat{P}\bar{y}$. Similarly, $V(\bar{y}) - \hat{Q}\bar{y} > V(\bar{x}) - \hat{Q}\bar{x}$. Adding up these two strict inequalities, $(\hat{P} - \hat{Q})(\bar{y} - \bar{x}) > 0$.

But, since $(P^{s'} - Q^{s'})(y^{s'} - x^{s'}) \to 0$ as $s \to \infty$, $(\hat{P} - \hat{Q})(\bar{y} - \bar{x}) = 0$, is a contradiction, which establishes the result.

References

Adler, R.L., Konheim, A.G., McAndrew, M.H., 1965. Topological entropy, Transactions of the American Mathematical Society 114, 309–319.

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

Block, L., Guckenheimer, J., Misiurewicz, M., Young, L.S., 1980. Periodic points and topological entropy of one dimensional maps. In: Nitecki, Z., Robinson, C. (Eds.), Global Theory of Dynamical Systems, Lecture Notes in Mathematics, 819, Springer-Verlag, Berlin, pp. 18–34.

Block, L.S., Coppel, W.A., 1992. Dynamics in One Dimension, Lecture Notes in Mathematics 1513, Springer-Verlag, Berlin.

Boldrin, M., Montrucchio, L., 1986. On the indeterminacy of capital accumulation paths, Journal of Economic Theory 40, 26–39.

Bowen, R., 1971. Entropy for group endomorphisms and homogeneous spaces, Transactions of the American Mathematical Society 153, 401–414.

Bowen, R., 1971. Periodic points and measures for Axiom A diffeomorphisms, Transactions of the American Mathematical Society 154, 377–397.

Day, R.H., 1994. Complex Economic Dynamics, Vol. 1, M.I.T. Press, Cambridge, MA.

Devaney, R.L., 1989. An Introduction to Chaotic Dynamical Systems, 2nd edn., Addison-Wesley, New York.

Dinaburg, E.I., 1970. The relation between topological entropy and metric entropy, Soviet Mathematics Doklady 11, 13–16.

Dutta, P.K., Mitra, T., 1989. Maximum theorems for convex structures with an application to the theory of optimal intertemporal allocation, Journal of Mathematical Economics 18, 77–86.

Gale, D., 1967. On optimal development in a multi-sector economy, Review of Economic Studies 34, 1–18.

Goodman, T.N.T., 1971. Relating topological entropy and measure entropy, The Bulletin of the London Mathematical Society 3, 176–180.

Goodwyn, L.W., 1969. Topological entropy bounds measure-theoretic entropy, Proceedings of the American Mathematical Society 23, 679–688.

McKenzie, L.W., 1986. Optimal economic growth, turnpike theorems and comparative dynamics. In: Arrow, K.J., Intriligator, M.D. (Eds.), Handbook of Mathematical Economics, Vol. III, North-Holland, Amsterdam.

Mitra, T., 1996. An exact discount factor restriction for period three cycles in dynamic optimization models, Journal of Economic Theory 69, 281–305.

Montrucchio, L., 1994. Dynamic complexity of optimal paths and discount factors for strongly concave problems, Journal of Optimization Theory and Applications 80, 385–406.

Montrucchio, L., Sorger, G., 1996. Topological entropy of policy functions in concave dynamic optimization models, Journal of Mathematical Economics 25, 181–194.

Nishimura, K., Yano, M., 1996. On the least upper bound of discount factors that are compatible with optimal period-three cycles, Journal of Economic Theory 69, 306–333.

Nishimura, K., Sorger, G., Yano, M., 1994. Ergodic chaos in optimal growth models with low discount rates, Economic Theory 4, 705–717.

Radner, R., 1961. Paths of economic growth that are optimal with regard only to final states, Review of Economic Studies 28, 98–104.

Sorger, G., 1992. On the minimum rate of impatience for complicated optimal growth paths, Journal of Economic Theory 12, 11–30.

Sorger, G., 1992b. Period three implies heavy discounting, mimeo.

Sorger, G., 1994. Period three implies heavy discounting, Mathematics of Operations Research 19, 1–16.

Stokey, N.L., Lucas, R.E., Jr., Prescott, E.C., 1989. Recursive Methods in Economic Dynamics, Harvard University Press, Cambridge.

Weitzman, M.L., 1973. Duality theory for infinite horizon convex models, Management Science 19, 783–789.