

ON DYNAMIC PROGRAMMING WITH UNBOUNDED RETURNS¹

*Jorge DURAN**

n° 9710

May 20, 1997. This version : June 4, 1997

Departamento de Economía, Universidad Carlos III de Madrid, Getafe, 28903 Madrid. Spain

and

IRES, Université Catholique de Louvain, Place Montesquieu 3, 1348-B Louvain-La-Neuve. Belgium

¹ This is part of my PhD dissertation. Helpful suggestions and comments are gratefully acknowledged to Aldo Rustichini and my advisor Manuel S. Santos. This paper was partly written while visiting the CEPREMAP at Paris with HCM program grant n° ERBCHRXCT940658. Financial support from HCM program grant n° ERBCHBGCT940681 is acknowledged. Comments to xurxo@eco.uc3m.es are welcome.

Jorge Duran

ON DYNAMIC PROGRAMMING WITH UNBOUNDED RETURNS

Abstract :

Some economic models like those of endogenous growth motivate the analysis of a class of recursive models sharing the property that the return function is not bounded along feasible paths. We consider a strategy of proof which allows to deal with many unbounded recursive models exploiting bounds to the rates of growth rather than to the levels.

Keywords : Dynamic Programming, Recursive Preferences, Unbounded Returns, Contraction Mappings, Weighted Norms.

PROGRAMMATION DYNAMIQUE AVEC UNE FONCTION DE RENDEMENT NON BORNEE

Résumé :

Différents modèles économiques, notamment de croissance endogène, justifient l'analyse d'une classe de modèles récurrents qui ont en commun la propriété que la fonction de rendement n'est pas bornée le long de sentiers praticables. Nous considérons une stratégie de démonstration qui s'applique à de nombreux modèles récurrents non bornés en exploitant des restrictions sur les taux de croissance plutôt que sur le niveau des variables considérées.

Mots-clés : Programmation dynamique, préférences récursives, rendements non bornés, applications contractantes, normes pondérées.

JEL classification : C61, D92

1. Introduction

In the analysis of recursive problems, the strongest and most general results are obtained in the case in which the return function is bounded along feasible paths. The key step using boundedness is to prove that the maximizing operator associated to Bellman equation has a fixed point which is afterwards proven to be the value function. Many economic models do not have bounded returns. Hence a weaker notion of boundedness must be considered. While exploring conditions under which an aggregator could determine a unique utility function, Boyd (1990) faces a similar problem; he tries to obtain a fixed point for a recursion operator when the aggregator is not bounded. This author proposes a weighted norm inducing a topology in the space of functions that is coarser than that induced by the supremum norm. It is in this weighted norm that he obtains the contraction property for the recursion operator.

In this paper we go a step further and see how these weighted topologies can be also applied to obtain a contraction for the maximizing operator when returns are unbounded. Most of the properties Boyd (1990) finds for the recursion operator are also inherited by the maximizing operator in some related weight function. It should be noted that all assumptions will be made on one-stage behavior of the elements of the model since it is dynamic programming tools what we want to develop. Also we consider the case of returns are unbounded from below and see how the argument does not apply anymore. The simple AK model is used to illustrate some key points of the argument.

The rest of this section presents the paper informally. Section 2 describes the model and reproduces the standard bounded case for non linear aggregators while section 3 proposes a generalization of these results while section 4 deals briefly with the M -contraction case. A last section discusses the limitations of this strategy of proof and relates this work to the literature.

1.1. The principle of optimality

Consider a dynamic program in which the programmer is endowed at each stage t with some predetermined state variable x_t and has to choose next stage's state x_{t+1} within a set of feasible choices $\Gamma(x_t)$ and a current action a_t from a set of admissible ones $\Omega(x_t, x_{t+1})$. With this specification, at each stage t it is clear that all the programmer cares about is the amount x_t with which is endowed. At some stage t and from a given initial condition x_t let us denote by $\Sigma(x_t)$ the collection of all feasible sequences of actions. Dynamic programming develops tools in order to deal with a class of problems like

$$\max U(a_t, a_{t+1}, \dots) \text{ s.t. } (a_t, a_{t+1}, \dots) \in \Sigma(x_t)$$

in which U has a very particular structure. A total return function like U is said to be recursive if there is a function W aggregating returns yielded by current actions and total returns from future actions

$$U(a_t, a_{t+1}, \dots) = W(a_t, U(a_{t+1}, a_{t+2}, \dots))$$

and this for all stage t . Thus, recursivity amounts to impose weak separability between current action and total discounted returns to future actions. Then, at each stage, the programmer is only concerned with current state x_{t+1} and action a_t as well as future choices for both variables as far as they induce a next stage's total return $U(a_{t+1}, a_{t+2}, \dots)$.

The principle of optimality first stated by Richard Bellman is based on the intuition that an optimal action can be characterized if we get concentrated only on the current action to be chosen and then let the programmer do the best she can with the state chosen for next stage. Define the value function $v(x_t)$ to be the value of U at the optimum in the problem above. Note that in essence, that problem starting in t is identical to one starting at $t + 1$. Hence, if at $t + 1$ we are given some x_{t+1} we should expect its optimal value to be given by $v(x_{t+1})$. The idea of "leaving the programmer do the best she can" may be formalized by substituting $U(a_{t+1}, a_{t+2}, \dots)$ by $v(x_{t+1})$ in the problem above while this is possible as soon as the axiom of weak separability between current action and future returns is imposed. The principle of optimality expects an optimal plan to be generated in some way as

$$v(x_t) = \sup_{x_{t+1} \in \Gamma(x_t)} \sup_{a_t \in \Omega(x_t, x_{t+1})} W(a_t, v(x_{t+1}))$$

for all t . This is the Bellman equation associated with our program. The analysis of this problem is much easier than the original one because an infinitely stages program has been reduced to a one step program.

Of course, this is done at the cost of dealing now with a functional equation whose unknown v will in general be impossible to obtain explicitly. A second important problem is that while the original program is expressed in terms of the total return function U the Bellman equation has only the aggregator W involved. Thus, the identification between these two functions is also an important part of the analysis.

1.2. Discounting the future

Lucas and Stokey (1984) proposed to use the aggregator as a primitive on which all assumptions were made and from which a return function was derived. This was done with the assumption that one-stage returns were in some sense bounded along admissible paths. Boyd (1990) examines under what assumptions on the aggregator a return function existed without assuming boundedness but bounding the rate of growth of returns. The idea of this paper is that those assumptions

identifying U and W also ensure that Bellman equation has a solution v with the necessary properties so as to ensure that the principle of optimality holds.

The basic idea is that W possesses the property that future returns are discounted in the sense that any increase in future return yields a proportionally smaller increase in current return. We require a number $0 < \delta < 1$ to exist such that given some action a and some level of future returns λ we have $W(a, \lambda) \leq W(a, 0) + \delta\lambda$ where $W(a, 0)$ can be interpreted as the return yielded by a single action and δ as an upper bound to the implicit discount factor. If this is true then $U(a_0, a_1, \dots) \leq \sum_{t=0}^{\infty} \delta^t W(a_t, 0)$ so what we need is that $W(a_t, 0)$ cannot be growing along any admissible path at a rate higher than that at which δ^t is converging to zero. In the bounded case one-stage return is assumed to be bounded and therefore discounting of δ^t must dominate from some period on. The idea of Boyd (1990) is to bound the rate of growth of returns. Since this may not be easy to do dealing directly with W this author seeks for a weight function: a function φ bounding in absolute value one-stage returns but with the additional property that $\varphi(a_t)$ cannot grow asymptotically at a factor greater than δ^{-1} . Hence, discounting eventually dominates the growth of returns along any admissible path. What we will show here is that this kind of transversality condition will also ensure that a solution to Bellman equation exists.

2. Non linear bounded aggregators

In this section we present the standard bounded case. We describe more formally the model we will work with and provide a condensed version of section 4.2 in Stokey and Lucas (1989) for non linear aggregators. The whole argument relies on boundedness of one-stage returns together with strict discounting.

2.1. Description of the model

The state space is $X \subset \mathbb{R}^n$. A continuous feasibility correspondence is $\Gamma : X \rightarrow X$ will describe feasible state paths: an accumulation path is a sequence $\mathbf{x} = (x_{t+1})_{t=0}^{\infty}$ and is said to be feasible from some initial condition $x_0 \in X$ when $x_{t+1} \in \Gamma(x_t)$ for all $t \geq 0$. Let us denote by $\Pi(x_0)$ the collection of all feasible from x_0 state paths. Also use $H \subset X \times X$ to denote the graph of Γ .

Let $A \subset \mathbb{R}^m$ be the space of actions and $\Omega : H \rightarrow A$ be the continuous admissibility correspondence. An action plan is a sequence $\mathbf{a} = (a_t)_{t=0}^{\infty}$ and is said to be feasible from some initial condition $x_0 \in X$ when there is at least one

$\mathbf{x} \in \Pi(x_0)$ such that $a_t \in \Omega(x_t, x_{t+1})$ for all $t \geq 0$. Denote by $\Sigma(x_0)$ the collection of all feasible from x_0 action plans and by $\Sigma = \bigcup_{x_0 \in X} \Sigma(x_0)$ the collection of all plans that are feasible from some initial condition. Also let $D = \Omega(H) \subset A$ and note that $\Sigma \subset D^\infty$.

The objective function is an extended real-valued total return function U on Σ . It is assumed to be recursive in the sense that there is an aggregator $W : D \times U(\Sigma) \rightarrow U(\Sigma)$ for which $U(\mathbf{a}) = W(\pi\mathbf{a}, U(\sigma\mathbf{a}))$ for all $\mathbf{a} \in \Sigma$. Note that π is denoting the first coordinate projection function while σ is the shift operator. We assume $|U(\mathbf{a})| < \infty$ for at least some $\mathbf{a} \in \Sigma$ so that we do not lose generality if we assume that $U(\Sigma) \ni 0$. The value function is defined as

$$v(x_0) = \sup_{\mathbf{a} \in \Sigma(x_0)} U(\mathbf{a}) \quad (2.1)$$

for all $x_0 \in X$. The optimization problem at the right-hand side of this equality is the problem whose solutions we aim to analyze. A plan \mathbf{a} is said to be optimal from $x_0 \in X$ when $\mathbf{a} \in \Sigma(x_0)$ and $v(x_0) = U(\mathbf{a})$. A solution to Bellman equation can be seen as a fixed point of the maximizing operator defined over all real-valued functions f on X as

$$Tf(x) = \sup_{y \in \Gamma(x)} \sup_{a \in \Omega(x, y)} W(a, f(y)) \quad (2.2)$$

for all $x \in X$. Dynamic programming arguments rely on finding some equivalence between these two problems. In particular we would like the value function to solve Bellman equation $v = Tv$ and a feasible program (\mathbf{x}, \mathbf{a}) to be optimal if and only if it verifies $v(x_t) = W(a_t, v(x_{t+1}))$ for all $t \geq 0$; when this is true we say that the principle of optimality holds.

An immediate observation is that while the original problem (2.1) is stated in terms of the return function, the Bellman equation (2.2) is written in terms of the aggregator. Hence the identification between U and W constitutes an important part of the analysis. Consider the recursion operator R defined over all extended real-valued functions V on Σ as $RV(\mathbf{a}) = W(\pi\mathbf{a}, V(\sigma\mathbf{a}))$ for all $\mathbf{a} \in \Sigma$. Note that a fixed point for this operator is a recursive return function.

In order to prove that the principle of optimality holds we will show that under some conditions the recursion operator R is a contraction in a suitable space of functions and that the return function U is recursive in the sense that $U = RU$. This will ensure that partial sums of returns $R^N 0$ will converge pointwise to the return function U which in turn helps to prove that the value function v is indeed the only bounded continuous solution to Bellman equation. As soon as v solves Bellman equation it is not difficult to show that the principle holds.

2.2. Returns in the limit

Following Lucas and Stokey (1984) and Boyd (1990), instead of starting from a total return function, an aggregator will act as a primitive in our model. The aim is to show the very close relationship between the recursion operator and the maximizing operator.

Let Λ be some closed subset of the extended real line containing zero. Let our aggregator be a continuous function $W : D \times \Lambda \rightarrow \Lambda$ increasing in its second argument and with the property that $\lambda \in \Lambda$ and $|\lambda| < \infty$ implies $|W(a, \lambda)| < \infty$ for all $a \in D$. Define the total return function as

$$U(\mathbf{a}) = \lim_{N \rightarrow \infty} W(a_0, W(a_1, \dots, W(a_{N-1}, 0) \dots)) = \lim_{N \rightarrow \infty} R^N 0(\mathbf{a}) \quad (2.3)$$

for all $\mathbf{a} \in \Sigma$. A bound to the returns to a single action and some kind of discounting will act as a transversality condition ensuring not only that U is well defined and recursive but also that v solves Bellman equation and that the principle of optimality holds. Assume:

(A1) One action return is bounded $\|R0\| < \infty$.

(A2) W is Lipschitz continuous of constant $\delta < 1$ in its second argument.

The following three propositions summarize the standard results for the bounded non linear aggregator case. They are stated and proved for later reference. We first see what is the relationship between the return function and the recursion operator providing us the link that will help us proving afterwards that the solution to Bellman equation is in fact the value function.

Proposition 2.1. *Under (A1) and (A2) the return function U defined in (2.3) is the only bounded and continuous fixed point of the recursion operator R .*

Proof. Endow Σ with the relative product topology and let $B(\Sigma)$ be the collection of bounded continuous real-valued functions on Σ together with the supnorm. Under (A2) the recursion operator possesses the Lipschitz property: let V, L be two elements of $B(\Sigma)$ and \mathbf{a} be an action plan. then

$$\begin{aligned} |RV(\mathbf{a}) - RL(\mathbf{a})| &= |W(\pi\mathbf{a}, V(\sigma\mathbf{a})) - W(\pi\mathbf{a}, L(\sigma\mathbf{a}))| \leq \delta |V(\sigma\mathbf{a}) - L(\sigma\mathbf{a})| \\ &\leq \delta \|V - L\|. \end{aligned}$$

Take the supremum over all \mathbf{a} and get $\|RV - RL\| \leq \delta \|V - L\|$. Let $V \in B(\Sigma)$ and note that W was assumed to be continuous and that π, σ are continuous

functions when the product topology is considered on $B(\Sigma)$. Then RV can be seen as the composition of continuous functions W, π, σ and V and therefore continuous itself. Using the Lipschitz property of R we have $\|RV\| \leq \delta \|V\| + \|R0\| < \infty$ under (A1) so that $RV \in B(\Sigma)$. Then R is a contraction of modulus δ on $B(\Sigma)$ and as a consequence of the contraction mapping theorem there is a unique element in this space U^* such that $U^* = RU^*$ and $\|R^N 0 - U^*\| \rightarrow 0$ so that it must be the case that $U^* = U$. Q.E.D.

Recall that this proves that U defined in (2.3) is bounded, continuous in the product topology and recursive in the sense that $U = RU$ as soon as boundedness (A1) and discounting (A2) hold. However, the importance of this result is that $R^N 0 \rightarrow U$ pointwise as this will help us proving that the maximizing operator is a contraction and U at the value function is its fixed point.

Proposition 2.2. *Under (A1) and (A2) the value function v defined in (2.1) is the only bounded and continuous fixed point of the maximizing operator T .*

Proof. Let $B(X)$ be the collection of bounded, continuous real-valued functions on X . Under (A2) the maximizing operator possesses the Lipschitz property: let $f, h \in B(X)$ and note that

$$\begin{aligned} |Tf(x) - Th(x)| &\leq \sup_{y \in \Gamma(x)} \sup_{a \in \Omega(x,y)} |W(a, f(y)) - W(a, h(y))| \\ &\leq \sup_{y \in \Gamma(x)} \delta |f(y) - h(y)| \leq \delta \|f - h\|. \end{aligned}$$

Taking the supremum over X we get $\|Tf - Th\| \leq \delta \|f - h\|$. Just as before is easy to see that $\|Tf\| \leq \delta \|f\| + \|R0\| < \infty$ as soon as (A1) holds and $f \in B(X)$. Continuity of Tf follows from the continuity assumptions on Γ, Ω and W , the choice of f as a continuous function and the maximum theorem. Hence $Tf \in B(X)$ and the contraction mapping theorem yields existence of f^* unique up to elements in $B(X)$ with the property that $f^* = Tf^*$.

Now consider the fact that for any $x_0 \in X$ and $\mathbf{x} \in \Pi(x_0)$ we have $\delta^N |f^*(x_N)| \rightarrow 0$ because $0 \leq \delta^N |f^*(x_N)| \leq \delta^N \|f^*\| \rightarrow 0$. Also $R^N \lambda$ is Lipschitz continuous of constant δ^N in its last argument and $R^N 0 \rightarrow U$ pointwise. Using these two facts the proof that $f^* = v$ follows closely the argument in theorem 4.3 in Stokey and Lucas (1989). Q.E.D.

Hence v is in fact the only bounded continuous solution to Bellman equation and that $T^N 0 \rightarrow v$ uniformly. While the first fact helps us proving that the

principle of optimality holds the second is at the core of many algorithms designed to approximate numerically solutions to (2.1).

Proposition 2.3. *Under (A1) and (A2) the principle of optimality holds.*

Proof. Let $x_0 \in X$ and $\mathbf{a} \in \Sigma(x_0)$ with \mathbf{x} satisfy $v(x_t) = W(a_t, v(x_{t+1}))$ for all t , then we may write $v(x_0) = R^N v(x_N)(\mathbf{a}) \leq R^N 0(\mathbf{a}) + \delta^N |v(x_N)| \leq R^N 0(\mathbf{a}) + \delta^N \|v\|$ which in the limit yields $v(x_0) \leq U(\mathbf{a})$. Since by definition of v it is true that $v(x_0) \geq U(\mathbf{a})$ both inequalities imply that this plan is optimal. The reverse way is exactly as in theorem 4.4 in Stokey and Lucas (1989) using now the fact that the aggregator is increasing in its second argument. Q.E.D.

The advantage of dynamic programming arguments is precisely that the awful task of analyzing the solutions to (2.1) directly reduces now to the analysis of a simple problem like

$$v(x) = \sup_{y \in \Gamma(x)} \sup_{a \in \Omega(x,y)} W(a, v(y)).$$

Under boundedness and discounting v has been shown to be continuous while W was assumed to be so. Also Γ and Ω were assumed to be compact-valued. Then for all x there is in fact a choice for a and y that attains the supremum. Further, the maximum theorem ensures that there is an upper semicontinuous policy correspondence $G : X \rightarrow D \times X$ such that $(a, y) \in G(x)$ implies $v(x) = W(a, v(y))$. For any initial condition $x_0 \in X$ construct a plan choosing each period t a pair $(a_t, x_{t+1}) \in G(x_t)$ so that $v(x_t) = W(a_t, v(x_{t+1}))$. The previous proposition ensures that this plan is optimal.

Further, we may be interested some comparative dynamics analysis so that the properties of G would be of importance. But we should be sure that G is in fact generating every optimal plan. And this is the case, let $\mathbf{a} \in \Sigma(x_0)$ with \mathbf{x} be optimal, since the principle of optimality holds we know that $v(x_t) = W(a_t, v(x_{t+1}))$ for all t so that it must be the case that $(a_t, x_{t+1}) \in G(x_t)$ for all t . It cannot be the case that some plan is optimal but it is not generated by the policy correspondence.

2.3. The case of linear technology

As an example consider the one sector model of economic growth with linear technology. A planner is assumed to produce the good in period t using the good of the previous period through a linear production function $f(x) = Ax$ for $A > 0$.

Here x is interpreted to be capital stock and A the net of depreciation marginal product of capital in the productive sector. Thus the feasibility correspondence is $\Gamma(x) = [0, Ax]$ for all $x \geq 0$. An action is a level of consumption; admissible consumption choices are those not exceeding output minus gross investment. The admissibility correspondence would then be $\Omega(x, y) = [0, Ax - y]$ for all $x \geq 0$ and $0 \leq y \leq Ax$.

Consider the Uzawa aggregator $W : \mathbb{R}_+ \times \mathbb{R}_- \rightarrow \mathbb{R}_-$ defined as $W(a, \lambda) = (-1 + \lambda)e^{-u(a)}$ where $u : \mathbb{R}_+ \rightarrow \mathbb{R}_{++}$ is some increasing function. First we know that W is bounded in its first argument because $|W(a, \lambda)| \leq |-1 + \lambda|$ so that (A1) holds. Also (A2) holds because

$$|W(a, \lambda) - W(a, \mu)| = e^{-u(a)}|\lambda - \mu| \leq e^{-u(0)}|\lambda - \mu|$$

therefore W is Lipschitz continuous of constant $e^{-u(0)} < 1$ because $u(0) > 0$. By proposition 2.1 we know then that

$$U(\mathbf{a}) = \lim_{N \rightarrow \infty} R^N 0(\mathbf{a}) = - \sum_{t=0}^{\infty} e^{-\sum_{j=0}^t u(a_j)}$$

is the utility function. Note that U takes values between $-e^{-u(0)}(1 - e^{-u(0)})^{-1}$ and zero because $e^{-u(0)} < 1$. The importance of the Lipschitz condition, of discounting, is clear because if $u(0)$ fails to be greater than zero $U(\mathbf{0}) = -\infty$ even if the boundedness condition holds as it is here the case. If this happens $v(0) = -\infty$ so that the value function also fails to be bounded.

Results like those in chapter 4 in Stokey and Lucas (1989) will prove other properties of v like in this case concavity and differentiability. With those properties at hand the analysis of

$$v(x) = \sup_{0 \leq y \leq Ax} (-1 + v(y))e^{-u(Ax-y)}$$

becomes a rather easy task while the principle of optimality will ensure that we are indeed characterizing all the solutions to the original problem.

3. Weighted bounded returns

Let us consider an identical model in which the boundedness assumption is relaxed and the discounting assumption will have now to be related to admissible sustainable rates of growth of returns. In particular we will allow returns to grow

unboundedly but with an upper bound to the factor at which they may do so that is sufficiently discounted by the Lipschitz constant of the aggregator.

Non linear aggregators can be seen as generating return functions whose implicit discount factor is bounded from above by its Lipschitz constant. As soon as the aggregator acts as a primitive and a return function is defined as in (2.3) the return function will in some sense be bounded from above by an additive function obtained from the aggregator and the Lipschitz constant. Then if $\mathbf{a} \in \Sigma$ we may write

$$R^N 0(\mathbf{a}) \leq \sum_{t=0}^{N-1} \delta^t W(a_t, 0). \quad (3.1)$$

In the bounded case it is obvious that as soon as $W(\cdot, 0)$ is bounded, the return function is bounded from above by $\|W(\cdot, 0)\| (1 - \delta)^{-1}$. However, it may be the case that $W(\cdot, 0)$ is not bounded but cannot grow steadily along any feasible path at a factor higher than δ^{-1} in which case the return function will also be bounded in some sense.

From a technical point of view the problem is to find a different complete space of functions on which our operators possess the Lipschitz property so that the contraction mapping theorem applies as in the bounded case. Let $C(D)$ be the collection of continuous and real-valued functions on D and let $\varphi \in C(D)$ be any function such that $\varphi > 0$. For any function $u \in C(D)$ define its weighted norm as $\|u\|_\varphi = \sup_{a \in D} |u(a)/\varphi(a)|$. Then let $C_\varphi(D)$ denote the subset of functions in $C(D)$ that are φ -bounded endow it with the φ -norm. The key fact exploited by Boyd (1990) is that the operator $S : B(D) \rightarrow C_\varphi(D)$ defined as $Su = \varphi \cdot u$ is a distance-preserving isomorphism so that $C_\varphi(D)$ inherits the completeness of $B(D)$.

3.1. Bounding one-stage growth of returns

In this case (A1) will be replaced by the requirement that $W(\cdot, 0)$ is φ -bounded for some $\varphi \in C(D)$ with $\varphi > 0$ and with $\delta\varphi(a_{t-1})/\varphi(a_t) < 1$ asymptotically along any admissible action path. Note that φ can be seen as a sort of absolute-value upper bound to the one-period return. Hence, this new requirement will amount to impose that one-period return cannot grow steadily at a factor greater than δ^{-1} so that sustained growth of future returns is sufficiently discounted by the Lipschitz property of W . The link between growing returns and technology will be given by function $\psi : W \rightarrow \mathbb{R}$ with the property that $\varphi(a_t) \leq \psi(x_t)$ along a

feasible state-action plan and such that one-stage feasible growth of ψ is strictly less than δ^{-1} . Assume:

(B1) Exists $\varphi \in C(D)$ with $\varphi > 0$ such that $\|R0\|_{\varphi} < \infty$.

(B2) W is Lipschitz continuous of constant δ . There is some $\psi \in C(X)$ such that (\mathbf{a}, \mathbf{x}) feasible implies $\varphi(a_t) \leq \psi(x_t)$ for all t and exists $\alpha > 0$ with

$$\delta \sup_{y \in \Gamma(x)} \frac{\psi(y)}{\psi(x)} \leq \alpha < 1 \text{ for all } x \in X. \quad (3.2)$$

It should be noted that as soon as φ has been found, a rather natural candidate for ψ is given by $\psi(x) = \sup_{y \in \Gamma(x)} \sup_{a \in \Omega(x,y)} \varphi(a)$. Under the continuity assumption on the feasibility and admissibility correspondences ψ is a continuous function on X taking positive real values. Hence $C_{\psi}(X)$ is a complete metric space in its ψ -norm.

Hence, we only need to find some ψ bounding $W(\cdot, 0)$ in absolute value and being discounted by δ in order to find a ψ -bounded fixed point f^* for the maximizing operator. Analogously, we need to find some Ψ bounding $W(\pi \cdot, 0)$ and being discounted by δ in order to find a Ψ -bounded fixed point U^* for the recursion operator. However, considering the way in which the recursion operator was defined, in terms of W , it should be natural to think that once ψ has been found, some Ψ must exist with the required properties.

Indeed, if (B2) holds we can choose some constant $\delta < \varsigma < 1$ small enough so that $\beta = \alpha\varsigma\delta^{-1} < 1$. Then $\varsigma \sup_{y \in \Gamma(x)} \psi(y)/\psi(x) \leq \beta < 1$ and let $\Psi : \Sigma \rightarrow \Lambda$ be defined as $\Psi(\mathbf{a}) = \sum_{t=0}^{\infty} \varsigma^t \varphi(a_t)$. This function is positive real-valued: as soon as \mathbf{a} with \mathbf{x} is feasible for some initial condition we have $\Psi(\mathbf{a}) \leq \sum_{t=0}^{\infty} \varsigma^t \psi(x_t) < \infty$ because $\varsigma\psi(x_{t+1})/\psi(x_t) \leq \beta < 1$ so that the series is summable. This implies that

$$\delta \frac{\Psi(\sigma \mathbf{a})}{\Psi(\mathbf{a})} = \delta \frac{\Psi(\sigma \mathbf{a})}{\varphi(a_0) + \varsigma \Psi(\sigma \mathbf{a})} \leq \delta \frac{\Psi(\sigma \mathbf{a})}{\varsigma \Psi(\sigma \mathbf{a})} \leq \frac{\delta}{\varsigma} < 1$$

for all $\mathbf{a} \in \Sigma$. Consider the coarser topology on Σ with respect to which Ψ is continuous; choose the finer of this and the relative product topology or the one generated by both collections if they are not ordered. This topology makes Ψ, π and σ continuous functions. Let $C_{\Psi}(\Sigma)$ be the collection of Ψ -bounded continuous real-valued functions on Σ . The following is the Continuous Existence Theorem of Boyd (1990) applied to our case:

Proposition 3.1. *Under (B1) and (B2) the return function U defined in (2.3) is the only Ψ -bounded and continuous fixed point of the recursion operator R .*

Proof. Under (B2) the recursion operator possesses the Lipschitz property: let V, L be two elements of $C_\Psi(\Sigma)$ and $\mathbf{a} \in \Sigma$, then

$$\begin{aligned} |RV(\mathbf{a}) - RL(\mathbf{a})| &\leq \delta \frac{|V(\sigma\mathbf{a}) - L(\sigma\mathbf{a})| \Psi(\sigma\mathbf{a})}{\Psi(\sigma\mathbf{a}) \Psi(\mathbf{a})} \Psi(\mathbf{a}) \\ &\leq \frac{\delta}{\varsigma} \|V - L\|_\Psi \Psi(\mathbf{a}). \end{aligned}$$

Divide the whole expression by $\Psi(\mathbf{a})$ and take the supremum over all $\mathbf{a} \in \Sigma$ to get $\|RV - RL\|_\Psi \leq \delta\varsigma^{-1} \|V - L\|_\Psi$. Let $V \in C_\Psi(\Sigma)$ and note that W was assumed to be continuous and that π, σ are continuous functions on Σ , then RV can be seen as the composition of continuous functions W, π, σ and V and therefore continuous itself. Using the Lipschitz property of R we have $\|RV\|_\Psi \leq \delta\varsigma^{-1} \|V\|_\Psi + \|R0\|_\Psi$ but under (B1) it is true that $\|R0\|_\Psi \leq \|R0\|_\varphi < \infty$ so $RV \in C_\Psi(\Sigma)$. Then R is a contraction of modulus $\delta\varsigma^{-1}$ on $C_\Psi(\Sigma)$ and as a consequence of the contraction mapping theorem there is a unique element in this space U^* such that $U^* = RU^*$ and $\|R^N 0 - U^*\|_\Psi \rightarrow 0$ so that $U^* = U$ because weighted uniform convergence still implies pointwise convergence as $\Psi > 0$. Q.E.D.

Again this proves that U defined in (2.3) is Ψ -bounded, continuous in some topology equally or stronger than the relative product topology on Σ and recursive in the sense that $U = RU$ as soon as weighted boundedness (B1) and net discounting (B2) hold. Note that the particular topology with respect to which U is continuous will not be of interest for us because it is the principle of optimality that we will use to show existence and characterize solutions. In practice, as soon as φ has a particular functional form it should not be difficult to find some weighted topology on Σ making this function continuous as Boyd (1990) does in analyzing a Ramsey problem.

Just as in the bounded case, pointwise convergence $R^N 0 \rightarrow U$ will help us proving the next result. Now the maximizing operator is shown to be a weighted contraction and that the value function is its unique weighted bounded fixed point.

Proposition 3.2. *Under (B1) and (B2) the value function v defined in (2.1) is the only ψ -bounded and continuous fixed point of the maximizing operator T .*

Proof. Under (B2) the maximizing operator possesses the Lipschitz property: let $f, h \in C_\psi(X)$ and note that

$$|Tf(x) - Th(x)| \leq \sup_{y \in \Gamma(x)} \delta \frac{|f(y) - h(y)| \psi(y)}{\psi(y) \psi(x)} \psi(x) \leq \alpha \|f - h\|_\psi \psi(x).$$

Divide by $\psi(x)$ and take the supremum over X to get $\|Tf - Th\|_\psi \leq \alpha \|f - h\|_\psi$. Now consider the fact that $T0(x) \leq \|R0\|_\psi \psi(x)$ so that $\|T0\|_\psi \leq \|R0\|_\psi < \infty$ under (B1), hence, $\|Tf\|_\psi \leq \alpha \|f\|_\psi + \|T0\|_\psi < \infty$ as soon as $f \in C_\psi(X)$. Continuity of Tf follows from the continuity assumptions on Γ, Ω and W , the choice of f as a continuous function and the maximum theorem. Hence $Tf \in C_\psi(X)$ and the contraction mapping theorem yields existence of f^* unique up to elements in $C_\psi(X)$ with the property that $f^* = Tf^*$.

Now note that for any $x_0 \in X$ and $\mathbf{x} \in \Pi(x_0)$ it is true that $\delta^N |f^*(x_N)| \rightarrow 0$ because $0 \leq \delta^N |f^*(x_N)| \leq \delta^N \|f^*\|_\psi \psi(x_N) \rightarrow 0$ which in turn is a consequence of the fact that (3.2) holds so that the series is summable. Also $R^N \lambda$ is Lipschitz continuous of constant $(\delta\zeta^{-1})^N$ in its last argument and $R^N 0 \rightarrow U$ pointwise. Again these two facts and the argument in theorem 4.3 in Stokey and Lucas (1989) prove that $f^* = v$. Q.E.D.

With these two results and in regard that $\delta^N |f^*(x_N)| \rightarrow 0$ along any feasible state path, the proof that the principle of optimality holds under (B1) and (B2) is exactly as in proposition 2.3.

3.2. Linear technology with homogeneous returns

Consider the AK model described in the previous section but now consider a linear aggregator $W : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined as $W(a, \lambda) = a^\theta + \delta \lambda$ where $\theta, \lambda \in (0, 1)$. Choose $\varphi(a) = 1 + a^\theta$ and note that in this case we can set $\psi(x) = 1 + (Ax)^\theta$. Clearly (B1) holds because φ is continuous and positive-valued on \mathbb{R}_+ . The aggregator W is Lipschitz continuous of constant δ in its second argument. Also note that

$$\delta \sup_{0 \leq y \leq Ax} \frac{1 + (Ay)^\theta}{1 + (Ax)^\theta} = \delta \frac{1 + (A^2x)^\theta}{1 + (Ax)^\theta} \leq \delta \lim_{x \rightarrow \infty} \frac{1 + (A^2x)^\theta}{1 + (Ax)^\theta} = \delta A^\theta$$

so that (B1) and (B2) will hold as soon as $\delta A^\theta < 1$. This simply requires the maximum rate of growth of returns to be strictly less than the inverse of the discount factor so that we ensure that growth is sufficiently discounted along any feasible path. In this particular case it should be noted that $\delta A^\theta < 1$ is not only a sufficient but also a necessary condition for the principle of optimality to hold; otherwise there will be an incentive to choose pure accumulation paths in order to consume infinite in the limit.

The observation that everything we care about ψ is that bounds φ from above was not spurious. Consider the same linear aggregator as above but now let

$\Gamma(x) = [0, Cx^\alpha + Ax]$ for some $C > 0$, $\alpha \in (0, 1)$ and $A > 1$. Since $\alpha < 1$, the only sustainable factor of growth of feasible accumulation paths is A because the marginal of the decreasing returns part of the production function tends to disappear as x grows. Intuitively $\delta A^\theta < 1$ should again be the necessary condition for (B1) and (B2) to hold. However, if we try to construct ψ as before we would have to deal with $1 + (Cx^\alpha + Ax)^\theta$ which is an awful function to work with in many applications. Instead, we may observe that there must be some constant $H > 0$ for which $Cx^\alpha + Ax \leq H + Ax$ for all $x \in X$. As this is true we can simply set $\psi(x) = 1 + (H + Ax)^\theta$ and check that

$$\begin{aligned} \delta \sup_{0 \leq y \leq Cx^\alpha + Ax} \frac{1 + (H + Ay)^\theta}{1 + (H + Ax)^\theta} &\leq \delta \frac{1 + (H + AH + A^2x)^\theta}{1 + (H + Ax)^\theta} \\ &\leq \delta \lim_{x \rightarrow \infty} \frac{1 + (H + AH + A^2x)^\theta}{1 + (H + Ax)^\theta} = \delta A^\theta. \end{aligned}$$

As we can see, again $\delta A^\theta < 1$ is all we need for the principle of optimality to hold in this model. This example also illustrates how asymptotic rates of growth is all we should care about. Scale effects like that induced by Cx^α do not matter as soon as they cannot be sustained in the long run.

4. M -stage contractions

When Denardo (1965) noted the contraction argument underlying dynamic programming, he already pointed out that in many applications an N -contraction rather than a contraction was all we could expect. In economic applications many technological specifications yield homogeneous feasible sets on which growth of returns can be bounded from above. However, it is often the case that some of these rates are not sustainable.

4.1. Discounting in many periods

Hence, sometimes, (3.2) can be a very strong assumption because it forces δ to discount rates of growth that are irrelevant in the long run. Sometimes a more careful choice of φ can help but some others this is not so straightforward. In those cases it may be easier to find an N -stage version of (3.2). Assume:

(C1) Exists $\varphi \in C(D)$ with $\varphi > 0$ such that $\|R0\|_\varphi < \infty$

(C2) W is Lipschitz continuous of constant δ . There is some $\psi \in C(X)$ such that (\mathbf{a}, \mathbf{x}) feasible implies $\varphi(a_t) \leq \psi(x_t)$ for all t and exists and exists $\alpha > 0$ and an integer $M \geq 1$ with

$$\delta^M \sup_{x_1 \in \Gamma(x_0)} \cdots \sup_{x_M \in \Gamma(x_{M-1})} \frac{\psi(x_M)}{\psi(x_0)} \leq \alpha < 1 \text{ for all } x_0 \in X. \quad (4.1)$$

Again, under (C2) is possible to find some $\delta < \varsigma < 1$ small enough so that $\beta = \alpha\varsigma^M\delta^{-M} < 1$ in which case (4.1) holds replacing δ by ς and α by β . If the rate of growth of ψ discounted by ς is bounded after M stages it must be case that the one-stage factor of growth is bounded so that there must be some $\gamma > 0$ such that $\varsigma \sup_{y \in \Gamma(x)} \psi(y)/\psi(x) \leq \gamma$. Now define Ψ as before and note that for any $\mathbf{a} \in \Sigma$ with \mathbf{x} associated

$$\begin{aligned} \Psi(\mathbf{a}) &\leq \sum_{t=0}^{\infty} \varsigma^t \psi(x_t) = \sum_{t=0}^{\infty} \sum_{j=0}^{M-1} \varsigma^{t+j} \psi(x_{t+j}) = \sum_{t=0}^{\infty} \varsigma^{tM} \psi(x_{tM}) \sum_{j=0}^{M-1} \varsigma^j \frac{\psi(x_{t+j})}{\psi(x_t)} \\ &\leq (1 + \gamma + \cdots + \gamma^{M-1}) \sum_{t=0}^{\infty} \varsigma^{tM} \psi(x_{tM}) \end{aligned}$$

which is a finite number because $\varsigma\psi(x_{(t+1)M})/\psi(x_{tM}) \leq \beta < 1$ so that the series is again summable. Using this weight function the statement and proof of proposition 3.1 would go without change. Hence R is still a contraction of modulus $\delta\varsigma^{-1}$ and U the only Ψ -bounded fixed point of the recursion operator. Of course, pointwise convergence still holds so that $R^N 0 \rightarrow U$.

The maximizing operator will now be an M -stage contraction of modulus α . The proof of proposition 3.2 goes without change except for the observation that now, for any $f, h \in C_\psi(X)$ we have

$$\begin{aligned} |T^M f(x_0) - T^M h(x_0)| &\leq \delta \sup_{x_1 \in \Gamma(x_0)} |T^{M-1} f(x_1) - T^{M-1} h(x_1)| \\ &\leq \delta^2 \sup_{x_1 \in \Gamma(x_0)} \sup_{x_2 \in \Gamma(x_1)} |T^{M-2} f(x_2) - T^{M-2} h(x_2)| \end{aligned}$$

so iterating up to M stages we get

$$\begin{aligned} |T^M f(x_0) - T^M h(x_0)| &\leq \delta^M \sup_{x_1 \in \Gamma(x_0)} \cdots \sup_{x_n \in \Gamma(x_{n-1})} \frac{|f(x_M) - h(x_M)| \psi(x_M)}{\psi(x_M) \psi(x_0)} \psi(x_0) \\ &\leq \delta^M \|f - h\|_\psi \sup_{x_1 \in \Gamma(x_0)} \cdots \sup_{x_M \in \Gamma(x_{M-1})} \frac{\psi(x_M)}{\psi(x_0)} \psi(x_0). \end{aligned}$$

but under (4.1) this inequality can be written as $|T^M f(x_0) - T^M h(x_0)| \leq \beta \|f - h\|_\psi \psi(x_0)$; divide by $\psi(x_0)$ and take the supremum over all $x_0 \in X$ to obtain $\|T^M f - T^M h\|_\psi \leq \beta \|f - h\|_\psi$. As T^M is a contraction of modulus $\alpha < 1$ there will be a unique f^* in $C_\psi(X)$ such that $f^* = T^M f^*$. But this will also be the unique fixed point for T in that space [see theorem 3.12 in Stokey and Lucas (1989)].

As a consequence, under (C1) and (C2) we can prove analogous results to those of propositions 3.1 and 3.2 so that the principle of optimality will again hold.

4.2. A model with human capital

Consider the standard model with physical and human capital. At each period t the planner is endowed with a certain amount k_t of physical capital and h_t of human capital. Some amount $0 \leq b_t \leq h_t$ of human capital is devoted to the physical good sector while the rest $h_t - b_t$ is invested in the human capital sector. Next period's state feasible choices are given by $0 \leq h_{t+1} \leq B h_t^\beta (h_t - b_t)^{1-\beta}$ where $B > 1$ and $0 < \beta < 1$ and by $0 \leq k_{t+1} \leq A k_t^\alpha b_t^{1-\alpha}$ for $A > 0$ and $0 < \alpha < 1$. Admissible consumption choices will then be given by $0 \leq c_t \leq A k_t^\alpha b_t^{1-\alpha} - k_{t+1}$. Let preferences be generated by $W(c, \lambda) = c^\theta + \delta \lambda$ where $\theta, \delta \in (0, 1)$.

It is not difficult to see that this technological specification yields continuous feasibility and admissibility correspondences. Note that the norm $\|k, h\| = \max\{|k|, |h|\}$ induces on $X = \mathbb{R}_+^2$ the usual metric. Note that for any given state $k, h \geq 0$ we have $k' \leq A k^\alpha h^{1-\alpha} \leq A \|k, h\|$ while $h' \leq B h \leq B \|k, h\|$. Choose the bigger of A and B and denote it by C . Then $(k', h') \in \Gamma(k, h)$ implies $\|k', h'\| \leq C \|k, h\|$. As in the linear case choose φ on $D = \mathbb{R}_+$ to be defined as $\varphi(c) = 1 + c^\theta$ and recall that any admissible from $k, h \geq 0$ consumption choice must verify $c \leq A k^\alpha h^{1-\alpha} \leq A \|k, h\| \leq C \|k, h\|$. Then choose $\psi(k, h) = 1 + C^\theta \|k, h\|^\theta$. As technology is homogeneous we found an upper bound C to the factor of growth of the norm of the state. However, this C may be influenced by scale effects that are irrelevant in the long run because it is only B that be need to discount.

Indeed, B is an upper bound to the rate of growth of human capital and therefore of the economy as the physical good sector is of constants returns to scale. Intuitively $\delta B^\theta < 1$ is all we need for (C1) and (C2) to hold. Consider a non feasible accumulation path constructed as $h_{t+1} = B h_t$ and $k_{t+1} = A k_t^\alpha h_t^{1-\alpha}$ for all t . Clearly, this is an upper bound to any feasible path but also if $x_t = k_t/h_t$ we have $x_{t+1} = A x_t^\alpha$ so that for N big enough $x_t, t \geq N$ must remain constant which

in turn implies that k_t is also growing at a factor B . Hence, for any $n \geq N$ we have $\|k_n, h_n\| \leq B \|k_{n-1}, h_{n-1}\|$ and therefore, for any initial condition $k_0, h_0 \geq 0$ and $n \geq N$ we have $\|k_n, h_n\| \leq B^{n-N} C^N \|k_0, h_0\|$. Then

$$\delta^n \frac{1 + C^\theta \|k_n, h_n\|^\theta}{1 + C^\theta \|k_0, h_0\|^\theta} \leq \delta^n \frac{1 + B^{\theta(n-N)} C^{\theta(N+1)} \|k_0, h_0\|^\theta}{1 + C^\theta \|k_0, h_0\|^\theta} \leq (\delta B^\theta)^{n-N} (\delta C^\theta)^N$$

so that as soon as $\delta B^\theta < 1$ we can choose some $M > N$ for which $(\delta B^\theta)^{M-N} (\delta C^\theta)^N < 1$. If this is the case (C1) holds for φ while (C2) holds for ψ and M . Then $\delta B^\theta < 1$ ensures that the principle of optimality holds for this problem.

5. Some comments

We have considered a slight variation of the old contraction argument to prove existence of a unique fixed point for a maximizing operator. By considering a weight function we relax the usual boundedness assumption by the requirement that an upper bound to the discount factor discounts the rates of growth of this weight function. One of the most disappointing aspects of this strategy of proof is that it does not provide with a general proof for those cases in which returns are unbounded from below.

5.1. Returns unbounded from below

Suppose that $\Lambda \subset \mathbb{R} \cup \{-\infty\}$ including $-\infty$ and that $W : D \times \Lambda \rightarrow \Lambda$ is continuous in the interior of its domain. However, let us now leave open the possibility that $W(a, \lambda) = -\infty$ for any $|\lambda| < \infty$ for some $a \in \partial D$. This is typically the case in any growth model, bounded from above or not, in which the value of zero consumption is $-\infty$. To obtain a contraction using the kind of φ used in the previous section is not possible unless very strong assumptions are made on the behavior of W on its domain. The difficulty here is due to the fact that actions implying a return of $-\infty$ are usually feasible. Thus, for $W(\cdot, 0)$ to be φ -bounded, the function φ must be chosen so that it takes unbounded values from above on the boundary of its domain.

Consider the case of logarithmic preferences $W(a, \lambda) = \log(a) + \delta\lambda$ for some $0 < \delta < 1$ in the context of the AK technology when $\Gamma(x) = [0, Ax]$ for all $x \in X = \mathbb{R}_{++}$ so that now $D = \mathbb{R}_{++}$. In this case the weight function may be $\varphi(a) = 1 + |\log(a)|$ which indeed is positive-valued, continuous on D and weights the one-stage return return function. However, as 0 is always an admissible action,

any admissible action plan \mathbf{a} can be such that $a_t = 0$ for some t in which case $\varphi(a_t) = \varphi(0) = \infty$ so that it is impossible to find a real-valued ψ function with the property that $\varphi(a_t) \leq \psi(x_t)$ for all t and all feasible state-action plan. Hence, the mere application of the previous argument to this case will not work unless we find some uniform lower bound to the rate of convergence to zero of feasible paths. It does not seem that this problem is solved by choosing φ more carefully.

Other authors have sorted this problem in various ways but none of them entirely satisfactory. In the Upper Semicontinuous Existence theorem, Boyd (1990) finds a function Ψ for which the recursion operator at zero is Ψ -bounded. However, he also assumes that $\delta \|\Psi \circ \sigma\|_\Psi < 1$; the example above shows that in practice it may not be easy to find a Ψ with such property.

5.2. Homogeneous programs

In a recent paper Alvarez and Stokey (1995) deal with a wide class of homogeneous problems. They work with the standard linear aggregator with discount factor $0 < \delta < 1$ and with an homogeneous of degree θ one-stage return function $F(x, y)$ defined on H . In their work they assume that $0 \notin X$ but also that $F : H \rightarrow \mathbb{R}$ and they bound the rates at which returns may be converging to $-\infty$ by assuming, for example in the case of $\theta < 0$, that there are positive constants b, B such that

$$0 < b \|x\|^\theta \leq |F(x, y)| \leq B (\|x\| + \|y\|)^\theta \text{ for all } (x, y) \in H.$$

This is a strong requirement, though. In the AK model with $X = \mathbb{R}_{++}$, $\Gamma(x) = [0, Ax]$ with $W(a, \lambda) = -a^\theta + \delta\lambda$ and $\theta < 0$ we have that $(x, Ax) \in H$ but $|F(x, Ax)| = |-0^\theta| = \infty$. That is, the standard AK model with returns unbounded from below does not meet such a requirement.

The point is that these author's assumption represents a bound to the rate of convergence to zero of feasible paths. When such a bound is imposed, a contraction may be obtained in many cases. Consider again the AK model with logarithmic one-period utility we saw in the previous subsection but now let $\Gamma(x) = [Jx, Ax]$ for some $0 < J < A$. A contraction is again obtained. Choose again $\varphi(a) = 1 + |\log(a)|$ and note that $\|R0\|_\varphi = 1$ so that (B1) holds. But also $|\log(a)|$ will be less or equal than $|\log(J) + \log(x)|$ or $|\log(A) + \log(x)|$ so that in any case $|\log(a)| \leq |\log(J)| + |\log(A)| + |\log(x)|$. Then choose $\psi(x) = 1 + |\log(J)| + |\log(A)| + |\log(x)|$ and note that it verifies (B2) because $\varphi(a_t) \leq \psi(x_t)$ along any feasible plan. Now simply note that

$$\delta \sup_{ax \leq y \leq Ax} \frac{\psi(y)}{\psi(x)} = \delta \frac{1 + |\log(J)| + |\log(A)| + \max\{|\log(J)|, |\log(A)|\} + |\log(x)|}{1 + |\log(J)| + |\log(A)| + |\log(x)|} \leq \delta < 1.$$

What happens is that the logarithm linearizes any exponential factor of divergence or convergence to zero so that as soon as $\delta < 1$ everything is working. But this depends crucially on the fact that $J > 0$ so that $|\log(J)| < \infty$. This same model with $W(a, \lambda) = -a^\theta + \delta\lambda$ and $\theta < 0$ would meet (B1) and (B2) as soon as $\delta J^{1-\sigma} < 1$. Again choose $\varphi(a) = 1 + a^\theta$ while $\psi(x) = 1 + (Jx)^\theta$. Indeed, in this case we should not care of returns growing from above but of convergence to $-\infty$ so that the principle of optimality will hold when δ discounts the factor J of convergence to zero of feasible paths.

5.3. Biconvergence

Streufert (1990, 1992) introduces the notion of biconvergence; a limiting condition ensuring that returns to any feasible path are on one hand sufficiently discounted from above (upper convergence) and from below (lower convergence). However, in those cases in which a one-stage return of $-\infty$ is admissible, the return function fails to be lower convergent. Also, as uniform discounting is substituted by a pointwise limit condition, he loses information about computational errors when iterating on the maximizing operator; he obtains convergence to the value function but pointwise convergence. His approach is interesting, though, as it allows to deal with models in which Lipschitz discounting is not possible to be obtained.

6. References

- Alvarez, Fernando and Nancy Stokey (1995) "Dynamic Programming with Homogeneous Functions." mimeo.
- Boyd III, John H. (1989) "Recursive Utility and the Ramsey Problem," *Journal of Economic Theory*, 50, 326-345.
- Denardo, Eric V. (1967) "Contraction Mappings in the Theory Underlying Dynamic Programming." *SIAM Review*, 9(2), 165-177.
- Lucas Jr., Robert E. and Nancy Stokey (1984) "Optimal Growth with Many Consumers," *Journal of Economic Theory*, 32, 139-171.
- Stokey, Nancy and Robert E. Lucas with Edward C. Prescott (1989) *Recursive Methods in Economic Dynamics*. Harvard University Press.
- Streufert, Peter A. (1990) "Stationary Recursive Utility and Dynamic Programming under the Assumption of Biconvergence," *Review of Economic Studies*, 57, 79-97.

Streufert, Peter A. (1992) "An Abstract Topological Approach to Dynamic Programming," *Journal of Mathematical Economics*, 21, 59-88.