DISCOUNTING LONG RUN AVERAGE GROWTH IN STOCHASTIC DYNAMIC PROGRAMS

Jorge Durán**

WP-AD 2002-08

Correspondence to: Jorge Durán, Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, Campus de San Vicente, 03080 Alicante, Spain, e-mail: duran@merlin.fae.ua.es.

Editor: Instituto Valenciano de Inve stigaciones Económicas, S.A.
First Edition July 2002.
Depósito Legal: V-2750-2002

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

---

* I am specially grateful to Cuong Le Van for fruitful discussions, encouragement, and for detecting an error in a previous version of this paper. Comments and suggestions are also acknowledged to Michele Boldrin, Raouf Boucekkine, Fabrice Collard, Tim Kehoe, Omar Licandro, and Luis Puch. I am also indebted to participants to the III Summer School on Economic Theory held at the Universidade de Vigo, the Macroeconomics Workshop at the Universitat Autònoma de Barcelona, and the Econometrics Seminar at Tilburg University. Financial support from PAI P4/01 project from the Belgian government at the IRES-UCL, from a European Marie Curie fellowship (Grant HPMF-CT-1999-00410) at the CEPREMAP, and from the IVIE is gratefully acknowledged.

** University of Alicante.

DISCOUNTING LONG RUN AVERAGE GROWTH IN STOCHASTIC DYNAMIC PROGRAMS

Jorge Durán

ABSTRACT

Finding solutions to the Bellman equation often relies on restrictive boundedness assumptions. In this paper we develop a method of proof that allows to dispense with the assumption that returns are bounded from above. In applications our assumptions only imply that long run average (expected) growth is sufficiently discounted, in sharp contrast with classical assumptions either absolutely bounding growth or bounding each period (instead of long run) maximum (instead of average) growth. We discuss our work in relation to the literature and provide several example

Keywords: Dynamic programming, Weighted norms, Contraction mappings, Dominated convergence, Non additive recursive functions.

JEL classification numbers: C61.

1 INTRODUCTION

The purpose of this paper is to provide easy-to-check conditions under which stochastic recursive models are represented by a well defined recursive optimization problem in turn characterized by its associated Bellman equation. Returns are assumed to be bounded from below but might be unbounded from above. Thus, our method of proof accounts for many economic models in which returns are not necessarily bounded from above: either because sustained growth is feasible or because some shock can affect unboundedly one period returns or both.

In the classical approach to recursive dynamic programming the value function is shown to solve the Bellman equation by proving first that the equation has indeed a solution. In order to do so, a maximization operator is defined (see Blackwell (1965) and Denardo (1967)) whose fixed points are solutions to the Bellman equation. This operator is then proven to be a contraction on the space of bounded continuous functions, a complete metric space in its supremum norm (in which case existence follows from the Banach fixed point theorem). The domain of the maximization operator can be seen as the class of admissible functions: candidates to solve the Bellman equation and, ultimately, to be the value function. However, the value function will only be bounded when returns to feasible choices are bounded and future returns are strictly discounted. Endogenous growth theory or business cycle theory provide numerous examples in which returns are not bounded. To overcome this problem several alternatives have been proposed. Stokey and Lucas's (1989, section 4.4) propose to work in spaces of homogeneous functions (with the norm over the unit circle) when dealing with deterministic homogeneous programs. This line of research has been followed by Álvarez and Stokey (1998) and Nakajima (1999) who also propose an approach to homogenous programs with returns unbounded from below. Streufert (1990, 1991) introduced the notion of biconvergence requiring only asymptotic discounting. His algorithm substitutes therefore uniform convergence (in some norm, underlying the contraction argument) by pointwise converge. His work is extended to the uncertain case in Streufert (1996), an analysis of a stochastic Ramsey model, and in a more general approach in Ozaki and Streufert (1996).

In the present paper we present a general stochastic recursive program in which returns are not necessarily bounded from above. Rather, average (expected) long run growth is required to be discounted. Returns are not assumed to be bounded but with respect to some weight function: it is this function's feasible growth that we require to be strictly discounted. Following Wessels (1977) we shall exploit the fact that spaces of weighted bounded functions are also complete. Hence, our maximization operator will still be a contraction on a complete metric space but in a larger class of functions (not necessarily bounded). As a result, a broad family of economic models can be reduced to a Bellman equation whose analysis can be carried out without awkward boundedness assumptions. Since the early work of Wessels (1977) the deterministic literature has shown an interest in functions that are only bounded with respect to some weight function. See Boyd (1990) who applies this technique to prove existence of a recursive utility function when the aggregator is not bounded or Dana and Le Van (1991) or Durán (2000) for applications to the analysis of deterministic recursive dynamic programs.

Our main accomplishment is the observation that, under a hypothesis easy to check in applications, weighted bounded functions are integrable and their integral continuous. Hence, admissible functions will be weighted bounded functions: with well defined expectation and continuous as functions of the endogenous state of the system, e.g., capital stock. If feasible growth of the weight function is strictly discounted the maximization operator will be shown to be a contraction on the class of weighted bounded functions. Further, once a solution to the Bellman equation has been found, a dominated convergence argument will allow us to reproduce familiar results in deterministic dynamic programming in order to ensure the connection between the original program and the Bellman equation. With these results at hand, our task in applications is to find a weight function and check whether its expectation is well behaved and its expected growth along feasible paths is sufficiently discounted. Our work partially generalizes Álvarez and Stokey (1998) on homogeneous dynamic programming and is related to Streufert (1996) and Ozaki and Streufert (1996) who analyze the case of asymptotic discounting of expected utility.

The main results are stated and proved in the next section. Section 3 describes the original program and shows how optimal paths can be generated by the policy correspondence. Our assumptions are discussed and our work examined in the context of the literature in section 4 where some limitations of this strategy are also discussed and future research suggested.

2 DISCOUNTING AVERAGE GROWTH

Some Borel subset $Z \subset \mathbb{R}^s$ acts as the exogenous state space: each period a random shock (e.g., a technological shock) is drawn from Z according to some Borel probability measure μ. Another Borel subset $X \subset \mathbb{R}^n$ acts as the endogenous state space (e.g., predetermined state variables as capital stock). Given some state (z, x), feasible endogenous state choices are described by $\Gamma(z, x)$ where $\Gamma : Z \times X \to X$ is a compact valued continuous correspondence.[1] A current state (z, x) and a feasible choice $y \in \Gamma(z, x)$ determine feasible actions (e.g., consumption) described as $c \in \Omega(z, x, y)$. Let H denote the graph of Γ, then $\Omega : H \to \mathbb{R}^m$ is compact valued and continuous while $\Omega(H)$ denotes its range. Following Lucas and Stokey (1984) the objective function (e.g., utility) will be constructed from an aggregator W treated as a primitive concept. Given an action c and a future return λ, current discounted return is given by $W(c, \lambda)$. Becker and Boyd (1997, chapters 1 and 3) constitutes an excellent motivation for the study of non additive objective functions.[2]

Example 1 Time additive objective functions are generated by the class of additive aggregators $W(c, \lambda) = u(c) + \delta\lambda$ where u is the one period reward function and $\delta \in (0, 1)$ the discount factor. An example of non additive aggregator is Uzawa's

[1] Hereafter the default topology in a product space is the product topology. Measurability will always refer to measurability with respect to the corresponding Borel σ algebra. Product spaces will be endowed with the product σ algebra unless otherwise stated. Many spaces involved in the results below are separable: on Cartesian products of separable Borel spaces the Borel and product σ algebras coincide (Lindelöf theorem).

[2] The present paper was motivated by the reading of Boyd (1990): an analysis of the conditions under which an unbounded aggregator function uniquely determines a total recursive utility function that is only bounded with respect to some weight.

$W(c, \lambda) = (-1 + \lambda)e^{-u(c)}$ where u is some increasing function with $u(0) > 0$. An unbounded non additive aggregator is $W(c, \lambda) = \log(\eta + c + \lambda)$ where $\eta > 1$.

The range of expected discounted returns is $\Lambda \subset \mathbb{R} \cup \{-\infty\}$ assumed to be closed and to contain zero. Assuming $\Lambda \ni 0$ only precludes the uninteresting case in which $-\infty$ is always the return to any action. $W : \Omega(H) \times \Lambda \to \Lambda$ is assumed to be continuous. Describing a stochastic recursive program in this paper means making explicit the nature of Γ, Ω, W, and μ. Alternative versions of the linear Ramsey or AK model shall be used throughout the paper to illustrate a number of points: see McGrattan (1998) for an overview of modern interpretations of the model.

Example 2 Consider the AK model of growth with random marginal product of capital. For stock of capital $x \geq 0$ and technological state $z \geq 0$, available choices for gross investment are $\Gamma(z, x) = [0, zx]$ while $\Omega(z, x, y) = [0, zx - y]$ are feasible consumption choices. Constant elasticity of intertemporal substitution preferences are generated by $W(c, \lambda) = c^\theta + \delta\lambda$ where $\delta \in (0, 1)$ and $0 < \theta \leq 1$.

From W we construct the total return function: the recursive program maximizes this function over feasible contingent plans, described in terms of Γ and Ω. We want the value function v to solve the Bellman equation

$$v(z, x) = \sup_{y \in \Gamma(z,x)} \sup_{c \in \Omega(z,x,y)} W(c, \int v(s, y)\, \mu(ds)). \tag{1}$$

Observe that we do not obtain the expectation of the aggregation but aggregate current actions with future expected returns. Kreps and Porteus (1978) show that temporal consistency of the underlying preferences imply that the aggregator is defined on expected future returns. Conversely, it is not difficult to see that having W defined on expected returns (the integral inside W) generates an objective function for which optimal plans are time-consistent. In order to prove that v verifies (1) we shall show first that the equation has at least a solution (proven afterwards to be the value function). Define the maximization operator M as

$$(Mf)(z, x) = \sup_{y \in \Gamma(z,x)} \sup_{c \in \Omega(z,x,y)} W(c, \int f(s, y)\, \mu(ds)). \tag{2}$$

where f is any candidate to be the value function. Solutions to the Bellman equation are the fixed points of the maximization operator. Our first task is to look for a class of admissible functions: candidates to be a solution to the Bellman equation and, ultimately, to be the value function. The class of admissible functions will be the domain of M. There are two reasonable requirements a function must meet in order to be admissible:

(a) An admissible function f must be integrable in the sense that the integral inside W in (2) is well defined. We should be able to compute the expectation of the value of any feasible choice.

(b) Even if the integral is well defined, in general it will only be upper semicontinuous as a function of y (Fatou theorem). Berge theorem applied to the optimization problem in (2), however, requires continuity. Álvarez and Stokey (1998) stand out the lack of a version of this theorem for upper semicontinuous functions: without continuity, existence of an optimal choice is not ensured.

Measurability will always be ensured by our continuity assumptions so that the (potential) third problem is avoided in this paper.[3] In short, we have to look for a class of functions for which expressions like $\int f(s, y)\, \mu(ds)$ make sense. This paper is on finding (displaying) a function, rather than a constant, that will bound candidates to solve the Bellman equation. Let g and φ be two continuous functions $\Omega(H) \to \mathbb{R}$ with $\varphi \geq 0$. We say that g is bounded with respect to φ when $\|g\|_\varphi = \|g/\varphi\| < \infty$ where $\|\,.\,\|$ denotes the supremum norm.

Assumption 1 There exists $\varphi : \Omega(H) \to \mathbb{R}_+$ continuous with $\|W(\,.\,, 0)\|_\varphi < \infty$.

Hence, there is a continuous function that absolutely bounds one period returns. In general, however, actions are free variables that can display very irregular behavior along feasible paths. We need to link this function to the state of the system; we do so assuming existence of some function of the state ψ bounding φ from above at feasible choices.

[3]See Stokey and Lucas (1989, page 388) for Blackwell's (1965) well known example illustrating the measurability problem. Bertsekas and Shreve (1978) is a discussion on the problem of measurability in dynamic programming. Most economic applications, however, meet our continuity assumptions.

Assumption 2 There exists $\psi : Z \times X \to \Lambda$ continuous with $\psi \geq 0$ such that $\varphi(c) \leq \psi(z,x)$ for all $(z,x) \in Z \times X$ and $c \in \Omega(z,x,y)$, some $y \in \Gamma(z,x)$.

This function will link potential feasible growth rates of returns with discounting and will bound admissible functions. The next assumption requires ψ to meet the two requirements (a) and (b) above; recall that in applications ψ is a displayed object whose properties are open to direct verification.

Assumption 3 $\int \psi(s,y)\,\mu(ds)$ is well defined and continuous in y.

Admissible functions will be chosen to be ψ-bounded functions because they share this continuous expectation property with their weight (and hence meet requirements (a) and (b) above). Observe that the supremum norm is a particular case of weighted norm when the weight function is chosen to be a constant. The bounded case automatically verifies assumptions 1 to 3 choosing $\varphi = \psi = 1$. The Uzawa aggregator is an example of this case. When $W(\,.\,,0)$ is not bounded φ and ψ cannot be chosen to be constant.

Example 3 In example 2, for all $c \geq 0$ we can simply choose $\varphi(c) = c^\theta$. Then let $\psi(z,x) = (zx)^\theta$. Assumptions 1 and 2 are verified by construction. Jensen's inequality implies $0 < \int (zx)^\theta\,\mu(dz) \leq (\bar{z}x)^\theta$ where $\bar{z} = \int z\,\mu(dz)$. Continuity follows from the Lebesgue dominated convergence theorem: assumption 3 holds.

The φ and ψ functions are not unique: other functions satisfying assumptions 1 to 3 exist but none can be constant because otherwise $W(c,0) = c^\theta > \varphi(c)$ at some c thus violating assumption 1.

The class of admissible functions will be that of ψ-bounded functions; if M is to be well defined on this set we should ensure that their expectation is well defined and continuous. An adaption of the proof of the Lebesgue dominated convergence theorem (Doob (1994, page 83)) allows us to prove an important result.

Lemma 1 *Under assumption 3 every ψ-bounded continuous function $f : Z \times X \to \mathbb{R}$ has the property that $\int f(z,x)\,\mu(dz)$ is well defined and continuous as a function of x.*

Proof: Suppose that $|f| \leq \psi$, otherwise follow the argument below for $\psi + \|f\|_\psi$ instead of ψ. Continuity ensures measurability of $f(\,.\,,x)$ while $|\int f(z,x)\,\mu(dz)| \leq \int |f(z,x)|\,\mu(dz) \leq \int \psi(z,x)\,\mu(dz)$ so that $f(\,.\,,x)$ is integrable. To see continuity let $(x_n) \subset X$ with $x_n \to x^0 \in X$. Since $0 \leq \psi - f$, Fatou theorem implies

$$\liminf_{n\to\infty} \int \psi(z,x_n) - f(z,x_n)\,\mu(dz) \geq \int \psi(z,x^0) - f(z,x^0)\,\mu(dz)$$

so that

$$\liminf_{n\to\infty} - \int f(z,x_n)\,\mu(dz) \geq - \int f(z,x^0)\,\mu(dz)$$

because $\lim_{n\to\infty} \int \psi(z,x_n)\,\mu(dz) = \int \psi(z,x^0)\,\mu(dz)$ by hypothesis. Hence,

$$\limsup_{n\to\infty} \int f(z,x_n)\,\mu(dz) \leq \int f(z,x^0)\,\mu(dz).$$

To end the proof note that $|f| \leq \psi$ implies $-f \leq \psi$ as well. The above argument applies to $0 \leq \psi + f$ to obtain

$$\int f(z,x^0)\,\mu(dz) \leq \liminf_{n\to\infty} \int f(z,x_n)\,\mu(dz)$$

as was to be shown. ■

The class of ψ-bounded continuous real valued functions $C_\psi(Z\times X)$ is complete in its ψ norm.[4] In order to ensure that M is a contraction on the class of ψ-bounded functions we need to ensure that ψ cannot be growing too fast (in regard to discounting) along feasible paths.

[4]Let $C_\psi(Z\times X)$ be the class of all continuous functions $Z \times X \to \Lambda$ that are ψ-bounded. When $\psi > 0$ let $B(Z\times X)$ be the class of bounded continuous functions and define u on $B(Z\times X)$ as $u(f) = f\psi$. Then u is a distance preserving isomorphism between $B(Z\times X)$ and $C_\psi(Z\times X)$, the latter endowed with the ψ norm. As a consequence $C_\psi(Z\times X)$ inherits the completeness of $B(Z\times X)$. The assumption that $\psi > 0$ is often made but is not necessary: observe that ψ bounded functions vanish, by definition, at $\ker(\psi)$; then apply the argument above with the (closed) subset $B'(Z\times X)$ of functions that vanish where ψ does so, instead of $B(Z\times X)$.

Assumption 4 W is Lipschitz continuous of constant $\delta < 1$ and increasing in its second argument. Further,

$$\delta \sup_{y \in \Gamma(z,x)} \frac{\int \psi(s,y)\, \mu(ds)}{\psi(z,x)} \leq \alpha \tag{3}$$

for some $0 < \alpha < 1$ and uniformly over $Z \times X$.

Under this assumption expected growth of ψ along a feasible path cannot exceed the factor α. More importantly, weighted bounded functions will be shown (next section) not to grow on average at a factor bigger than α in the long run: they can grow faster for short periods but not sustainably. In some cases expected growth can always be discounted as in the following (homogeneous) program.

Example 4 Let ψ be as in example 3. Given (z, x) feasible choices y verify $\int (sy)^\theta \, \mu(ds) \leq y^\theta \int s^\theta \, \mu(ds) \leq (zx)^\theta \bar{z}^\theta$ where we use Jensen's inequality. Then

$$\delta \sup_{0 \leq y \leq zx} \frac{\int (sy)^\theta \, \mu(ds)}{(zx)^\theta} \leq \delta \bar{z}^\theta$$

so that assumption 4 holds as soon as $\delta \bar{z}^\theta < 1$: only expected feasible growth must be discounted.

Nevertheless, there are many cases in which feasible paths might display transitory high rates of growth of returns (and therefore value) while these are not sustainable in the long run. Streufert (1996) emphasizes that the literature on Lipschitz discounting has often required to always (every period) discount maximum (instead of expected) growth. Yet, weight functions allow, first, to discount only average growth (as example 4 illustrates); second, to care only for long run growth. Indeed, the weight function ψ acts as an intermediary between discounting and the admissible function: the weight can verify (3) every period while it weights functions for which this is only true in the long run (figure 1 is an illustration for functions of one variable). As a consequence, with a careful choice of the weight, we can account for non homogeneous programs in which average growth of returns is only discounted in the long run.

Jones and Manuelli (1990) describe a deterministic model with these characteristics. In their model the value function can be proven to exist, to solve Bellman

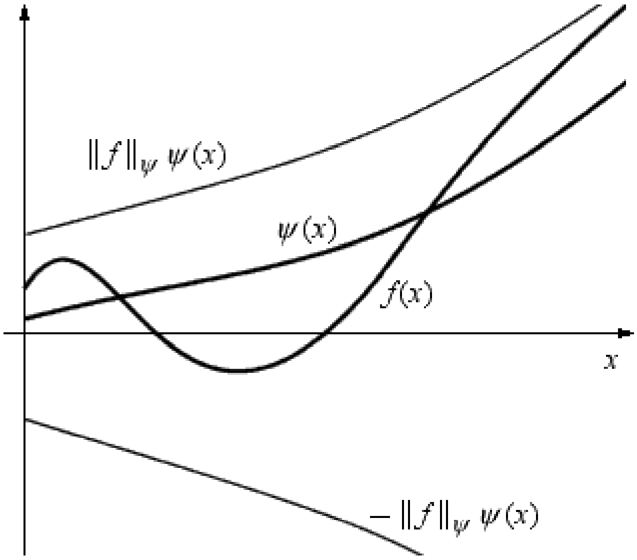

Figure 1: A function f that is ψ bounded

equation, and to be continuous finding a simple weight function: see Durán (2000, section 2.4). The following example constitutes further one possible extension of this model to the stochastic case.

Example 5 Modify example 2 to consider $\Gamma(z,x) = [0, zx + x^\rho]$ where $\rho \in (0,1)$. The graph of Γ is not a cone: close to the origin the term x^ρ allows for expected growth factors that are not sustainable. Nevertheless, let $\psi(z,x) = (\eta + zx + x^\rho)^\theta$ for some $\eta > 0$ and assumption 4 again requires $\delta \bar{z}^\theta < 1$. Indeed, use the change of variable $w = zx + x^\rho$ to show that

$$\delta \sup_{0 \le y \le zx + x^\rho} \frac{\int (\eta + sy + y^\rho)^\theta \, \mu(ds)}{(\eta + zx + x^\rho)^\theta} \le \delta \sup_{0 \le w} \left(\frac{\eta + \bar{z}w + w^\rho}{\eta + w} \right)^\theta .$$

The quotient in brackets is a continuous function of w valued 1 at zero and $\bar{z}$ when w is taken to infinity. Somewhere in between there is a maximum value. When $\delta \bar{z}^\theta < 1$ we can always choose η big enough so as to make this maximum value less than one in which case assumption 4 holds (when $\delta \bar{z}^\theta \ge 1$ we can always exceed one choosing w big enough no matter how big η was chosen).

Observe that the non-linear part of the production function could have been affected by some other shock as well. Underlying this example is the fact that many specifications induce feasible sets behaving asymptotically and in expectation as a cone. Weighted bounded functions help us abstracting from temporary effects like those induced by x^ρ in the example.

Assumptions 1 to 3 together with lemma 1 ensure that M is well defined on the class of ψ-bounded functions, our admissible functions. Assumption 4 will further ensure that M is a contraction on this class of functions.

Remark 1 To prove that M is a contraction Blackwell's (1965) sufficient conditions are not necessary if one notes that in general

$$|\max_x p(x) - \max_x q(x)| \leq \max_x |p(x) - q(x)|.$$

This observation together with Lispschitz continuity of W directly yield the Lipschitz property of M in the proof below.

There are versions of Blackwell's sufficient conditions for weighted norms (as in Boyd (1990)) but they are somewhat difficult to check. Further, the proof does not require W to be non decreasing in its second argument (although we do not use this fact in this paper because W will be increasing).

Proposition 1 *Under assumptions 1 to 4 the maximization operator M has a fixed point f^*, unique up to elements in $C_\psi(Z \times X)$.*

Proof: Under assumption 4 the maximizer operator has the Lipschitz property. Let $f, h \in C_\psi(Z \times X)$ and fix any $(z, x) \in Z \times X$. Then

$$\begin{aligned}
|(Mf)(z,x) - (Mh)(z,x)| &\leq \delta \sup_{y \in \Gamma(z,x)} |\int f(s,y) - h(s,y)\, \mu(ds)| \\
&\leq \delta \sup_{y \in \Gamma(z,x)} \int \frac{|f(s,y) - h(s,y)|}{\psi(s,y)} \frac{\psi(s,y)}{\psi(z,x)}\, \mu(ds)\psi(z,x) \\
&\leq \delta \sup_{y \in \Gamma(z,x)} \int \frac{\psi(s,y)}{\psi(z,x)}\, \mu(ds)\, \|f - h\|_\psi \, \psi(z,x) \\
&\leq \alpha \|f - h\|_\psi \, \psi(z,x),
\end{aligned}$$

where the first inequality follows from Lipschitz continuity of W in its second argument and the remark above, and the second to assumption 4. Dividing both sides by $\psi(z,x)$ and taking the supremum over $Z \times X$ yields $\|Mf - Mh\|_\psi \leq \alpha \|f - h\|_\psi$. Now note that for all $f \in C_\psi(Z \times X)$ it is true that

$$\|Mf\|_\psi \leq \alpha \|f\|_\psi + \|M0\|_\psi \leq \alpha \|f\|_\psi + \|W(\,.\,,0)\|_\varphi < \infty,$$

by the triangular inequality and because assumptions 1 and 2 imply $\|M0\|_\psi \leq \|W(\,.\,,0)\|_\varphi$ legitimating the second inequality. Hence, Mf is bounded in the ψ norm. Continuity follows from continuity of W, f, Γ and Ω, from the fact that both Γ and Ω take on compact values, from assumption 3, lemma 1, and Berge theorem. Then M maps $C_\psi(Z \times X)$ into itself and is a contraction. The proposition follows from the Banach fixed point theorem. ■

This proposition establishes existence of a (ψ-bounded) solution to the Bellman equation. The task left is to prove that v exists and that $v = f^*$. This is so under assumptions 1 to 4, assumed to hold throughout the next section.

In short, given some recursive program described by Γ, Ω, W, and μ we have to look for functions φ and ψ verifying assumptions 1 to 4. In such case (the next section proves that) the value function exists, solves Bellman equation, and is a continuous ψ-bounded function. Further, the policy correspondence generates optimal plans and every optimal plan is almost everywhere equal to a generated plan. The reader uninterested in the technical details can go directly to section 4 in which some additional examples serve as illustration of the advantages and shortcomings of this strategy of proof.

3 THE PRINCIPLE OF OPTIMALITY

In this section we will see that the total return function (defined below) is well defined. In general recursive preferences δ acts as an upper bound to the implicit discount factor. This is no surprise if one considers that (under our assumptions) feasible growth of $\varphi \leq \psi$ is strictly discounted by δ while actual returns are bounded by these functions and discounted at a factor at most equal to δ.

We first describe the original program. We denote $Z^t = Z \times \cdots \times Z$ (t times). A contingent state plan is a sequence $\mathbf{x} = (x_{t+1})_{t=0}^\infty$ of measurable functions $x_{t+1} : Z^t \to X$ and is feasible from (z_0, x_0) when $x_{t+1}(z^t) \in \Gamma(z_t, x_t(z^{t-1}))$ for all

z^t and $t \geq 0$. For any initial condition $\Pi(z_0, x_0)$ stands for the class of feasible plans. A contingent actions' plan is a sequence $\mathbf{c} = (c_t)_{t=0}^{\infty}$ of measurable functions $c_t : Z^t \to X$ and is feasible from (z_0, x_0) when $c_t(z^t) \in \Omega(z_t, x_t(z^{t-1}), x_{t+1}(z^t))$ for all z^t, all $t \geq 0$ and some $\mathbf{x} \in \Pi(z_0, x_0)$. We write $\Sigma(z_0, x_0)$ for the class of feasible actions' plans. Recall that $Z \times X$ is a Borel set and Γ compact valued and upper semicontinuous: theorem 7.6 in Stokey and Lucas (1989) ensure existence of a measurable selection from Γ so that Π is non empty. Since Γ has closed (and therefore Borel) graph an analogous argument ensures that Σ is non empty. For all $t \geq 1$ consider the family of functions $Z^t \to \Omega(H)$, each endowed with the distance of the maximum of one and the supremum of the difference. The class $\mathbf{C} = \bigcup_{(z,x)} \Sigma(z, x)$ is contained in the Cartesian product of these spaces: endow it with the relative product topology. For any pair $(z, \mathbf{c}) \in Z \times \mathbf{C}$ the continuation of $\mathbf{c}$ from z is a new contingent plan $\sigma(z, \mathbf{c})$ the tth coordinate $\sigma_t(z, \mathbf{c})$ defined to be the restriction of the $t + 1$st coordinate of $\mathbf{c}$ to $\{z\} \times Z^{t-1}$. Since restrictions of measurable functions to measurable sets are measurable we have $\sigma(z, \mathbf{c}) \in \mathbf{C}$.

The total return function is defined (in terms of W) as the pointwise limit of partial sums of returns. The recursion operator R maps every continuous function $V : \mathbf{C} \to \Lambda$ to a new function RV defined as

$$(RV)(\mathbf{c}) = W(\pi\mathbf{c}, \int V(\sigma(z, \mathbf{c}))\, \mu(dz)) \tag{4}$$

for all $\mathbf{c} \in \mathbf{C}$ where π denotes the first coordinate projection function. A function V is recursive when $V = RV$. The equation $V = RV$ is sometimes referred to as Koopmans equation. As in the case of the maximization operator, we will have to describe a class of admissible functions: candidates to be a fixed point V^* of R. The total return function U is defined at every $\mathbf{c} \in \mathbf{C}$ as

$$U(\mathbf{c}) = \lim_{N \to \infty} (R^N 0)(\mathbf{c}) \tag{5}$$

where R^N denotes the N times composition of R and 0 the constant function equal to zero. The value function is therefore defined as $v(z, x) = \sup_{\mathbf{c} \in \Sigma(z,x)} U(\mathbf{c})$ for all (z, x). A plan $\mathbf{c}$ is optimal for (z, x) when $\mathbf{c} \in \Sigma(z, x)$ and $v(z, x) = U(\mathbf{c})$. Recall that $\Sigma(z, x)$ is not empty so that v will be well defined as soon as U is so.

Observe that δ is an upper bound to the implicit discount factor and that φ bounds feasible returns. It is therefore reasonable to expect an additive function with less discounting $\varsigma > \delta$ and immediate returns φ to weight the total return function. Indeed, choose $\delta < \varsigma < 1$ small enough so that $\beta = \varsigma\alpha/\delta < 1$ (this is always possible under assumption 4). In such case (3) still holds with δ and α substituted by ς and β respectively. Define Φ on $\mathbf{C}$ as

$$\Phi(\mathbf{c}) = \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t(z^t))\, \mu^t(dz^t) \tag{6}$$

where μ^t denotes the product μ measure on Z^t. Under our assumptions this function is well defined, continuous, and the expression $\int \Phi(\sigma(z, \mathbf{c}))\, \mu(dz)$ is well defined and continuous as a function of $\mathbf{c}$ (see lemmas 4 and 6 in appendix A).

Lemma 2 *Every Φ bounded continuous function $V : \mathbf{C} \to \mathbb{R}$ has the property that $\int V(\sigma(z, \mathbf{c}))\, \mu(dz)$ is well defined and continuous as a function of $\mathbf{c}$.*

The proof follows an argument analogous to that of the proof of lemma 1 and is to be found in appendix A. This result is important because it implies that R is well defined on the class of Φ bounded continuous functions, our class of admissible functions. The total return function U happens to belong to this class: the following is a stochastic version of Boyd's (1990) continuous existence theorem where Φ has been displayed and its discounting checked.

Proposition 2 *The total return function is well defined and the unique fixed point of the recursion operator R in the space $C_\Phi(\mathbf{C})$.*

Proof: By construction, for any $\mathbf{c} \in \mathbf{C}$ it is true that

$$\delta \frac{\int \Phi(\sigma(z, \mathbf{c}))\, \mu(dz)}{\Phi(\mathbf{c})} = \delta \frac{\int \Phi(\sigma(z, \mathbf{c}))\, \mu(dz)}{\varphi(c_0) + \varsigma \int \Phi(\sigma(z, \mathbf{c}))\, \mu(dz)} \leq \frac{\delta}{\varsigma} < 1 \tag{7}$$

because $\varphi \geq 0$ and of the choice of ς. Then let $V, L \in C_\Phi(\mathbf{C})$ and fix $\mathbf{c} \in \mathbf{C}$. We have

$$\begin{aligned} |(RV)(\mathbf{c}) - (RL)(\mathbf{c})| &\leq \delta \int \frac{|V(\sigma(s, \mathbf{c})) - L(\sigma(s, \mathbf{c}))|}{\Phi(\sigma(s, \mathbf{c}))} \frac{\Phi(\sigma(s, \mathbf{c}))}{\Phi(\mathbf{c})}\, \mu(ds) \Phi(\mathbf{c}) \\ &\leq \frac{\delta}{\varsigma} \|V - L\|_\Phi\, \Phi(\mathbf{c}) \end{aligned}$$

where we have used Lipschitz continuity of W and (7). Divide by $\Phi(\mathbf{c})$ and take the supremum over $\mathbf{C}$ to obtain $\|RV - RL\|_{\Phi} \leq \delta\varsigma^{-1} \|V - L\|_{\Phi}$. Further, RV is continuous whenever $V \in C_{\Phi}(\mathbf{C})$ because it is a composition of continuous functions, by hypothesis and by lemma 2. Hence, R is a contraction of modulus $\delta\varsigma^{-1} < 1$ so that exists a unique V^* with $V^* = RV^*$ and $\left\|R^N 0 - V^*\right\|_{\Phi} \to 0$ (Banach fixed point theorem). Convergence in the Φ norm still implies pointwise convergence: it must be the case that $V^* = U$. ∎

In short, under our assumptions U is well defined and recursive with respect to W. To prove that the value function does solve the Bellman equation we shall also use an interesting property, pointed out informally in section 2, of weighted bounded functions under assumption 4.

Lemma 3 *Let f be ψ-bounded and continuous, $\mathbf{x}$ feasible, and assumption 4 hold. Then, we have*

(a) the function $f(z_t, x_t(z^{t-1}))$ is μ^t-integrable, and

(b) the series $\delta^t \int f(z_t, x_t(z^{t-1}))\, \mu^t(dz^t)$ is absolutely summable.

Proof: Let f be ψ-bounded and continuous and $\mathbf{x}$ be feasible from some initial condition (z_0, x_0). For any $t \geq 1$ we have

$$0 \leq \delta^t \left| \int f(z_t, x_t(z^{t-1}))\, \mu^t(dz^t) \right| \leq \|f\|_{\psi}\, \delta^t \int \psi(z_t, x_t(z^{t-1}))\, \mu^t(dz^t)$$

while the last term is finite. Indeed, under assumption 4 it is true that

$$\delta \int \psi(z_t, x_t(z^{t-1}))\, \mu(dz_t) \leq \alpha \psi(z_{t-1}, x_{t-1}(z^{t-2}))$$

for all z^{t-1}. Use this inequality and Fubini theorem to show that

$$\delta^t \int \psi(z_t, x_t(z^{t-1}))\, \mu^t(dz^t) \leq \alpha \delta^{t-1} \int \psi(z_{t-1}, x_{t-1}(z^{t-2}))\, \mu^{t-1}(dz^{t-1}) \leq \alpha^t \psi(z_0, x_0).$$

Hence, the expression $f(\,.\,, x_t(\,.\,))$ is μ^t integrable because $\alpha^t \psi(z_0, x_0) < \infty$. Moreover, since $\alpha < 1$ the series, both in terms of f and ψ, converges to zero as $t \to \infty$ and is absolutely summable because α^t is so. ∎

This result, together with recursivity of U stated above, will prove that the solution found to the Bellman equation is indeed the value function. Once the value function is shown to be ψ-bounded this lemma will again be used to prove that plans generated by the policy correspondence are optimal and viceversa. Summarizing:

Theorem 1 *Let assumptions 1 to 4 hold. The value function is well defined, continuous, and ψ-bounded. There is an optimal plan. If a plan $(\mathbf{c}, \mathbf{x})$ is feasible from (z_0, x_0) and verifies*

$$v(z_t, x_t(z^{t-1})) = W(c_t(z^t), \int v(z_{t+1}, x_{t+1}(z^t))\, \mu(dz_{t+1})) \tag{8}$$

for all z^t and $t \geq 0$, then it is optimal. Conversely, if a plan is optimal, this equality holds for μ^t almost every z^t and all $t \geq 0$.

The proof of this theorem makes use of the policy correspondence associated to the Bellman equation (1). The policy correspondence $G : Z \times X \to X \times \Omega(H)$ is defined as

$$\begin{aligned} G(z, x) = \{(y, c) \quad : \quad & y \in \Gamma(z, x),\ c \in \Omega(z, x, y) \\ & \text{and } v(z, x) = W(c, \int v(s, y)\, \mu(ds))\} \end{aligned} \tag{9}$$

for all (z, x). A plan $(\mathbf{c}, \mathbf{x})$ is said to be generated by the policy correspondence from some initial condition (z_0, x_0) when $(x_{t+1}(z^t), c_t(z^t)) \in G(z_t, x_t(z^{t-1}))$ for all z^t and $t \geq 0$. A generated plan is therefore a plan verifying (8). In appendix B corollary 2 proves that generated plans are optimal while corollary 3 shows that there is a generated (and therefore optimal) contingent plan. Finally proposition 4 shows that an optimal plan is equal to a generated plan μ^t almost everywhere.

Under the appropriate assumptions, v can be shown to possess other important properties such as monotonicity, concavity, or differentiability. The related results presented in Stokey and Lucas (1989, chapters 4 and 9) go with no change once U and v have been shown to be well defined, the first solving the Koopmans equation and the second the Bellman equation.

4 SOME COMMENTS

Describing the elements of our program we have done assumptions that deserve some comment. The aggregator W is a function of actions and not of the state of the system. This assumption was done for the sake of clarity of exposition and does no harm. If we need the state variables to directly affect current returns we can always define Ω so as to allow only one possible action, namely the required function of the state. Note that the notion of recursive function is precisely based on weak separability between current and future actions: modelling actions as we did does not imply any further structure on the underlying preference order.

Example 6 An agent inherits some savings $x \geq 0$ rewarded by a constant gross interest $R > 0$ and receives a random income $z \geq 0$, the remaining income being automatically consumed. Feasible next period's savings are described by $\Gamma(z, x) = [0, Rx + z]$ while $\Omega(z, x, y) = \{Rx + z - y\}$ instead of $[0, Rx + z - y]$. (If the agent has additive preferences as in the AK model of section 2, then $\delta R^\theta < 1$ is sufficient for (3) to hold for $\psi(z, x) = (\eta + Rx + z)^\theta$, some $\eta > 0$ big enough. Such condition does not depend on the particular mean of income z.)

Not so inoquous is the assumption that the random shock is independently and identically distributed (according to μ). It makes the analysis simpler, although at some cost of generality. Note, however, that many correlated cases can still be accounted for by the analysis above. Indeed, many correlated programs are described by an explicit stochastic law of motion in which some iid element is the ultimate source of uncertainty.

Example 7 Consider a correlated version of example 2: given a marginal product of capital $a \geq 0$ and a state of nature $z \geq 0$ current marginal product of capital is given by za^γ where $\gamma \in (0, 1)$ and $z \sim \mu$. The marginal product of capital is now an endogenous state variable. Then $\Gamma(z, a, x) = \{za^\gamma\} \times [0, za^\gamma x]$ while Ω and W are as before.

When facing a particular recursive program, the application of our results require displaying the weight functions; each case requires a specific weight. There is, however, no systematic procedure to construct these functions.

The lack of a systematic procedure to obtain suitable weights may induce two problems. First, a wrong choice of the weight functions can be misleading:

Example 8 Suppose $W(c, \lambda) = \log(\gamma + c + \lambda)$, when $\gamma > 1$ the aggregator is Lipschitz continuous of constant γ^{-1} in its second argument. Let Γ be as in the AK model. Assumption 1 holds for $\varphi(c) = \gamma + c$ but $\psi(z, x) = \gamma + zx$ yields $\gamma^{-1}\bar{z}^{\theta} < 1$ as a condition for (3) to hold. Choose instead $\varphi'(c) = \eta + \log(\gamma + c)$ and $\psi'(z, x) = \eta + \log(\gamma + zx)$ for some $\eta > 0$. We have

$$\frac{1}{\gamma} \sup_{0 \leq y \leq zx} \frac{\eta + \int \log(\gamma + sy)\, \mu(ds)}{\eta + \log(\gamma + zx)} \leq \frac{1}{\gamma} \frac{\eta + \log(\gamma + \bar{z}zx)}{\eta + \log(\gamma + zx)}.$$

This quotient is a continuous function of zx valued one both at zero and at the limit when $zx \to \infty$. An argument similar to that of example 5 shows that assumptions 1 to 4 hold as soon as $\gamma > 1$ (just choose η big enough).

In this example, the first φ did weight returns but too much; the weighted norm induced a topology much coarser than we needed. Indeed, first, the logarithm linearizes any exponential growing path and, second, such aggregator implies a very strong discounting pattern because as c grows, the implicit discount factor becomes smaller (zero in the limit). In other cases, the problem may be the reverse: some weight function might not work, while there is indeed a weight for which all our assumptions hold. In particular, the fact that some functions φ and ψ verify assumptions 1 to 3 does not imply that either assumption 4 holds for this functions or for no function at all.

Second, and more importantly, we can find cases in which finding the weight is far from obvious.

Example 9 In example 7 above we have $c^{\theta} \leq (za^{\gamma}x)^{\theta}$ so one may conclude too quickly that $\psi(z, a, x) = (za^{\gamma}x)^{\theta}$ is a good weight because assumptions 1 to 3 hold. Nevertheless, we have

$$\delta \sup_{(b,y) \in \Gamma(z,a,x)} \frac{\int (sb^{\gamma}y)^{\theta}\, \mu(ds)}{(za^{\gamma}x)^{\theta}} \leq \delta \left(\frac{\bar{z}(za^{\gamma})^{\gamma} za^{\gamma}x}{za^{\gamma}x} \right)^{\theta} = \delta \bar{z}^{\theta} (za^{\gamma})^{\gamma\theta},$$

a number that can be arbitrarily large for large values of z.

In example 7, a certain factor productivity the previous period has two effects: first, an immediate effect on output in that period and therefore in potential capital stock today; second, a lagged effect on today's factor productivity. In both cases, a higher factor productivity the previous period increases current output. Nevertheless, while $(za^{\gamma}x)^{\theta}$ in the definition of ψ in example 9 is able to account for the first effect, it does not account for the second, and hence the unbounded expression $(za^{\gamma})^{\gamma\theta}$ in maximum feasible growth.

This does not mean that the problem is not well defined; it just means that we have not chosen the correct weight. However, example 9 is a case in which a suitable weight could not be found (unless some restrictive assumption is made on the support of μ).

The most important shortcoming of the strategy developed above refers to returns unbounded from below. When describing our recursive program, the range of possible values for returns was defined as $\Lambda = \mathbb{R} \cup \{-\infty\}$. This definition, however, can be misleading as it can suggest that we can deal with programs with returns unbounded from below while this is not true in general. When feasible actions allow for an immediate return of $-\infty$ at any stage, it will not be possible to find a function like ψ in assumption 2.

Example 10 Consider the AK model of example 2 but with $W(c, \lambda) = \log(c) + \delta\lambda$. In this case assumption 1 could hold with $\varphi(c) = 1 + |\log(c)|$. Nevertheless, $0 \in \Gamma(z, x)$ for all $z, x \geq 0$ so that $\varphi(0) = \infty$ is always feasible. Hence, there is no real valued function ψ with the property that $\varphi \leq \psi$ for feasible choices.

Durán (2000) further discusses this example in the deterministic setting: introducing uncertainty leaves the problem unchanged. Álvarez and Stokey (1998) propose in the deterministic case an alternative interesting approach for homogeneous programs based on monotonicity.

Let us end this section recalling the flexibility of this approach to the analysis of the Bellman equation. Many economic theories are supported by models amenable to be formally expressed in terms of identical or analogous objects as the Bellman equation analyzed in this paper. A seminal paper by Loury (1981) analyzes the dynamics of inequality in an overlapping generations model economy with altruistic educational investment. Dutta and Michel (1998) consider a similar model where it is bequests rather than educational investment the mechanism of transmission of inequality across generations. In both cases, if the offspring's indirect utility function enters an agent's utility function, two consecutive generations's preferences are link by some Bellman-type functional equation.

Example 11 Consider a stochastic version of Loury's (1981) economy. Any agent is initially endowed with some stock of human capital h (also equal to income: effective labor productivity is equal to one). The agent decides consumption $c \geq 0$ and investment $e \geq 0$ in her children's education subject to $c + e \leq h$. Offspring human capital is $g(z, e)$ where $z \geq 0$ is some productivity shock with distribution μ. Given consumption c and offspring utility λ an agent derives utility $W(c, \lambda)$. Assume $g : \mathbb{R}_+^2 \to \mathbb{R}_+$ and $W : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ are continuous. An indirect utility function v is said to be consistent (across generations) when

$$v(h) = \sup_{0 \leq e \leq h} W(h - e, \int v(g(s, e)) \, \mu(ds))$$

for all $h \geq 0$. Suppose that W is Lipschitz continuous of constant $\delta < 1$ in its second argument and that there is some continuous $\psi : \mathbb{R}_+ \to \mathbb{R}$ with $\|W(\,.\,, 0)\|_\psi < \infty$ and

$$\delta \sup_{0 \leq e \leq h} \frac{\int \psi(g(s, e)) \, \mu(ds)}{\psi(h)} \leq \alpha < 1$$

for some $\alpha > 0$. The argument in proposition 1 is readily adapted to prove (using lemma 1) that a unique consistent indirect utility function v exists, that is continuous and ψ-bounded.

This literature requires fairly strong assumptions preventing the economy from growing in the long run (human capital is just a device to generate inequality). Those assumptions, however, in regard of our results, are unnecessary. Thus, such these interesting models can be integrated in the model economies used to assess the relationship between inequality and growth (see, for example, Aghion and Bolton (1997)) or the short run dynamics of inequality in growing economies (see the motivating paper by Díaz-Giménez et al (1997)).

A THE WEIGHT OF TOTAL RETURNS

Throughout the two appendixes, assumptions 1 to 4 are assumed to hold.

Lemma 4 *Function* Φ *described in (6) is well defined on* $\mathbf{C}$*, takes on finite values, and is continuous.*

Proof: Let $(\mathbf{c}, \mathbf{x})$ be a feasible plan. For any $t \geq 0$ the function $\varphi(c_t(\,.\,))$ is measurable: it is a composition of a measurable and a continuous function. Under assumption 2 we have $\varphi(c_t(z^t)) \leq \psi(z_t, x_t(z^{t-1}))$ for all z^t and hence lemma 3(a) ensures is μ^t integrable while

$$\Phi(\mathbf{c}) \leq \sum_{t=0}^{\infty} \varsigma^t \int \psi(z_t, x_t(z^{t-1}))\, \mu^t(dz^t).$$

Since (3) holds with δ and α replaced by ς and β, follow the same steps as in the proof of lemma 3(b) to prove that $\Phi(\mathbf{c})$ is finite.

To see continuity let $(\mathbf{c}^n) \subset \mathbf{C}$ with $\mathbf{c}^n \to \mathbf{c}^0 \in \mathbf{C}$ coordinatewise. The proof follows in three steps, each reproducing the strategy of lemma 1.

Step 1. Since ψ is continuous $\psi(z_t, x_t^n(z^{t-1})) \to \psi(z_t, x_t^0(z^{t-1}))$ for all z^t and all t. Under assumption 3 we have that

$$\lim_{n\to\infty} \int \psi(z_1, x_1^n)\, \mu(dz_1) = \int \psi(z_1, x_1^0)\, \mu(dz_1). \tag{10}$$

and

$$\lim_{n\to\infty} \int \psi(z_2, x_2^n(z_1))\, \mu(dz_2) = \int \psi(z_2, x_2^0(z_1))\, \mu(dz_2) \tag{11}$$

for all z_1. Under assumption 4 and given the way ς and β were chosen, we have that

$$\varsigma \int \psi(z_2, x_2^n(z_1))\, \mu(dz_2) \leq \beta\psi(z_1, x_1^n)$$

for all $n \in \mathbb{N}$ and all z_1, and hence

$$\beta\psi(z_1, x_1^n) - \varsigma \int \psi(z_2, x_2^n(z_1))\, \mu(dz_2) \geq 0.$$

Use continuity of ψ and while (11) to write

$$\lim_{n\to\infty}\left(\beta\psi(z_1,x_1^n)-\varsigma\int\psi(z_2,x_2^n(z_1))\,\mu(dz_2)\right)$$
$$=\beta\psi(z_1,x_1^0)-\varsigma\int\psi(z_2,x_2^0(z_1))\,\mu(dz_2)$$

for all z_1. In regard of these two expressions Fatou theorem implies

$$\liminf_{n\to\infty}\int\left(\beta\psi(z_1,x_1^n)-\varsigma\int\psi(z_2,x_2^n(z_1))\,\mu(dz_2)\right)\,\mu(dz_1)$$
$$\geq\beta\int\psi(z_1,x_1^0)\,\mu(dz_1)-\varsigma\int\int\psi(z_2,x_2^0(z_1))\,\mu(dz_2)\mu(dz_1).$$

Then use (10) to write

$$\liminf_{n\to\infty}-\varsigma\int\psi(z_2,x_2^n(z_1))\,\mu^2(dz^2)\geq-\varsigma\int\psi(z_2,x_2^0(z_1))\,\mu^2(dz^2),$$

where we have already applied Fubini theorem, and conclude that

$$\limsup_{n\to\infty}\int\psi(z_2,x_2^n(z_1))\,\mu^2(dz^2)\leq\int\psi(z_2,x_2^0(z_1))\,\mu^2(dz^2).$$

Since $\psi>0$ we can apply an analogous argument to ψ directly and write

$$\liminf_{n\to\infty}\int\psi(z_2,x_2^n(z_1))\,\mu^2(dz^2)\geq\int\psi(z_2,x_2^0(z_1))\,\mu^2(dz^2)$$

but these last two inequalities together imply

$$\lim_{n\to\infty}\int\psi(z_2,x_2^n(z_1))\,\mu^2(dz^2)=\int\psi(z_2,x_2^0(z_1))\,\mu^2(dz^2). \tag{12}$$

For the next step use (12) instead of (10) and

$$\lim_{n\to\infty}\int\psi(z_3,x_3^n(z^2))\,\mu(dz_3)=\int\psi(z_3,x_3^0(z^2))\,\mu(dz_3)$$

instead of (11) to prove the same result for $t=3$. Proceed recursively to show that

$$\lim_{n\to\infty}\int\psi(z_t,x_t^n(z^{t-1}))\,\mu^t(dz^t)=\int\psi(z_t,x_t^0(z^{t-1}))\,\mu^t(dz^t) \tag{13}$$

for all $t\geq0$.

Step 2. Under assumption 2 we have $\varphi(c_t^n(z^t)) \leq \psi(z_t, x_t^n(z^{t-1}))$ for all z^t and all t. Then $\psi(z_t, x_t^n(z^{t-1})) - \varphi(c_t^n(z^t)) \geq 0$ and an argument identical to that of step 1 yields

$$\lim_{n\to\infty} \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) = \int \varphi(c_t^0(z^t))\, \mu^t(dz^t) \tag{14}$$

for all $t \geq 0$.

Step 3. Now let us prove that the entire sum is also continuous. From the proof of lemma 3 and given the way ς and β where chosen we know that

$$\varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \leq \beta^t \psi(z_0, x_0^n)$$

for all $n \in \mathbb{N}$ and all t. Then

$$\beta^t \psi(z_0, x_0^n) - \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \geq 0$$

can be seen as a function of t integrable with respect to the counting measure. We will use again an argument similar to that of lemma 1: Fatou theorem implies

$$\liminf_{n\to\infty} \sum_{t=0}^{\infty} \left(\beta^t \psi(z_0, x_0^n) - \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t)\right)$$
$$\geq \sum_{t=0}^{\infty} \left(\beta^t \psi(z_0, x_0^0) - \varsigma^t \int \varphi(c_t^0(z^t))\, \mu^t(dz^t)\right)$$

where we have used continuity of ψ and (14) to solve the limits at the right hand-side of the inequality. Again because ψ is continuous

$$\liminf_{n\to\infty} \sum_{t=0}^{\infty} \beta^t \psi(z_0, x_0^n) = \liminf_{n\to\infty} \frac{\psi(z_0, x_0^n)}{1-\beta} = \frac{\psi(z_0, x_0^0)}{1-\beta}$$

so that the inequality above implies

$$\liminf_{n\to\infty} -\sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \geq -\sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^0(z^t))\, \mu^t(dz^t)$$

and finally

$$\limsup_{n\to\infty} \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \leq \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^0(z^t))\, \mu^t(dz^t).$$

Apply Fatou theorem directly to $\varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \geq 0$ and use again (14) to conclude that

$$\liminf_{n\to\infty} \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) \geq \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^0(z^t))\, \mu^t(dz^t).$$

These last two inequalities imply that

$$\lim_{n\to\infty} \Phi(\mathbf{c}^n) = \lim_{n\to\infty} \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^n(z^t))\, \mu^t(dz^t) = \sum_{t=0}^{\infty} \varsigma^t \int \varphi(c_t^0(z^t))\, \mu^t(dz^t) = \Phi(\mathbf{c}^0)$$

as was to be shown. ■

The expected value of the continuation of a contigent plan is continuous in the value of the shock considered, as stated in the following result.

Lemma 5 *Let* $V : \mathbf{C} \to \mathbb{R}$ *be a continuous function, then* $V(\sigma(z, \mathbf{c}))$ *is measurable as a function of* z *and continuous as a function of* $\mathbf{c}$.

Proof: Fix $z \in Z$ and let $(\mathbf{c}^n) \subset \mathbf{C}$ with $\mathbf{c}^n \to \mathbf{c}^0 \in \mathbf{C}$. Coordinatewise convergence implies

$$\begin{aligned} &\sup_{z^{t-1}\in Z^{t-1}} |\sigma_t(z, \mathbf{c}^n)(z^{t-1}) - \sigma_t(z, \mathbf{c}^0)(z^{t-1})| \\ &\qquad = \sup_{z^{t-1}\in Z^{t-1}} |c_{t+1}(z, z^{t-1}) - c_{t+1}(z, z^{t-1})| \leq \sup_{z^t\in Z^t} |c_{t+1}(z^t) - c_{t+1}(z^t)| \to 0 \end{aligned}$$

so that $\sigma(z, .)$ is continuous and therefore $V(\sigma(z, .))$ is the composition of continuous functions. Now fix $\mathbf{c} \in \mathbf{C}$ and consider a sequence of truncated plans: for all integer $k \geq 1$ write $\sigma(z, \mathbf{c})_k$ for a sequence with $\sigma_t(z, \mathbf{c})_k = \sigma_t(z, \mathbf{c})$ for all $0 \leq t \leq k$ and $\sigma_t(z, \mathbf{c})_k = c \in \Omega(H)$ for all $t > k$. Clearly $\sigma(z, \mathbf{c})_k \to \sigma(z, \mathbf{c})$ in the product topology as $k \to \infty$. Since V is continuous $V(\sigma(z, \mathbf{c})_k) \to V(\sigma(z, \mathbf{c}))$. Every $V(\sigma(z, \mathbf{c})_k)$ is a measurable function; $V(\sigma(z, \mathbf{c}))$ has been therefore shown to be the pointwise limit of a sequence of measurable functions. ■

Let $C_\Phi(\mathbf{C})$ be the class of functions $\mathbf{C} \to \Lambda$ continuous with respect to the relative product topology and bounded in norm Φ.

Lemma 6 *For all* $\mathbf{c} \in \mathbf{C}$ *the function* $\Phi(\sigma(\,.\,,\mathbf{c}))$ *is integrable. Further, the expression* $\int \Phi(\sigma(z,\mathbf{c}))\,\mu(dz)$ *is continuous as a function of* $\mathbf{c}$.

Proof: First note that lemma 4 establishes continuity of Φ while lemma 5 ensures measurability of $\Phi(\sigma(\,.\,,\mathbf{c}))$ as a function of z. Next observe that for any $\mathbf{c} \in \mathbf{C}$ with $\mathbf{x}$ associated and $z \in Z$ we have

$$\int \Phi(\sigma(z,\mathbf{c}))\,\mu(dz) = \int \left\{ \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \int \varphi(c_{t+1}(z,z^t))\,\mu^t(dz^t) \right\} \mu(dz)$$

where each partial sum inside the big brackets is a relative product measurable function. Note also that

$$\varsigma^t \int \varphi(c_{t+1}(z,z^t))\,\mu^t(dz^t) \le \varsigma^t \int \psi(z_t, x_{t+1}(z,z^t))\,\mu^t(dz^t) \le \beta^t \psi(z,x_1)$$

and therefore

$$0 \le \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \int \varphi(c_{t+1}(z,z^t))\,\mu^t(dz^t) \le \lim_{N\to\infty} \sum_{t=0}^{N} \beta^t \psi(z,x_1) = \frac{\psi(z,x_1)}{1-\beta}.$$

As a consequence for every $\mathbf{c} \in \mathbf{C}$ and $z \in Z$ partial sums always converge (to a finite real number). The limit function is measurable as it is the pointwise limit of a sequence of measurable functions. It is μ-integrable because it is dominated by $(1-\beta)^{-1}\psi(z,x_1)$, an integrable function under assumption 3.

We now prove continuity: let $(\mathbf{c}^n) \subset \mathbf{C}$ with $\mathbf{c}^n \to \mathbf{c}^0 \in \mathbf{C}$ and note that according to the expression above

$$\begin{aligned}
\lim_{n\to\infty} \int \Phi(\sigma(z,\mathbf{c}^n))\,\mu(dz) &= \lim_{n\to\infty} \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \int \left(\int \varphi(c^n_{t+1}(z,z^t))\,\mu^t(dz^t) \right) \mu(dz) \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \int \varphi(c^n_{t+1}(z^{t+1}))\,\mu^{t+1}(dz^{t+1}) \\
&= \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \lim_{n\to\infty} \int \varphi(c^n_{t+1}(z^{t+1}))\,\mu^{t+1}(dz^{t+1})
\end{aligned}$$

where the first equality follows from Beppo-Levi theorem, the second from Fubini theorem, and the third from lemma 4. In regard of (14) we have

$$\lim_{n\to\infty} \int \Phi(\sigma(z,\mathbf{c}^n))\,\mu(dz) = \lim_{N\to\infty} \sum_{t=0}^{N} \varsigma^t \int \varphi(c^0_{t+1}(z^{t+1}))\,\mu^{t+1}(dz^{t+1})$$

but again by Fubini and Beppo-Levi theorems respectively

$$\begin{aligned}\lim_{n\to\infty}\int \Phi(\sigma(z,\mathbf{c}^n))\,\mu(dz) &= \lim_{N\to\infty}\int\left\{\sum_{t=0}^{N}\varsigma^t\int\varphi(c^0_{t+1}(z,z^t))\,\mu^t(dz^t)\right\}\mu(dz)\\ &= \int\lim_{N\to\infty}\left\{\sum_{t=0}^{N}\varsigma^t\int\varphi(c^0_{t+1}(z,z^t))\,\mu^t(dz^t)\right\}\mu(dz)\\ &= \int\Phi(\sigma(z,\mathbf{c}^0))\,\mu(dz)\end{aligned}$$

as was to be shown. ■

With these results at hand we can prove lemma 2 ensuring that Φ bounded functions are reasonable admissible functions.

Proof of lemma 2: By lemma 5 we know that $V(\sigma(z,\mathbf{c}))$ is measurable as a function of z and continuous as a function of $\mathbf{c}$. Also V is bounded in norm Φ by hypothesis so that

$$\int V(\sigma(z,\mathbf{c}))\,\mu(dz)\leq\|V\|_\Phi\int\Phi(\sigma(z,\mathbf{c}))\,\mu(dz),$$

integrable by lemma 6. It is therefore well defined.

Let $(\mathbf{c}^n)\subset\mathbf{C}$ be such that $\mathbf{c}^n\to\mathbf{c}^0\in\mathbf{C}$. Since $V(\sigma(z,\mathbf{c}))$ is continuous in $\mathbf{c}$ we know that $V(\sigma(z,\mathbf{c}^n))\to V(\sigma(z,\mathbf{c}^0))$ pointwise. Since V is dominated by Φ and this last function has the property that $\int\Phi(\sigma(z,\mathbf{c}))\,\mu(dz)$ is continuous in $\mathbf{c}$ it follows, from an argument analogous to the proof of lemma 1, that $\int V(\sigma(z,\mathbf{c}))\,\mu(dz)$ is also continuous as a function of $\mathbf{c}$. ■

B THE PRINCIPLE OF OPTIMALITY

The proofs of this appendix are adapted from those in chapters 4 and 9 in Stokey and Lucas (1989) in order account for non additive aggregators and ψ-bounded value functions. Once we have ensured that U and v are well defined we will exploit the fact that $R^N 0\to U$ pointwise to prove that v actually solves the Bellman equation.

Proposition 3 *The value function v is well defined and the unique solution to the Bellman equation in the space $C_\psi(Z \times X)$.*

Proof: That is well defined follows from $\Sigma(z, x) \neq \emptyset$ for all (z, x) and from proposition 2 ensuring that U is well defined. Let (z_0, x_0) be any initial condition and $\mathbf{c} \in \Sigma(z_0, x_0)$ with $\mathbf{x}$ associated. Proposition 1 ensures existence of a solution f^* to the Bellman equation. Then

$$\begin{aligned} f^*(z_0, x_0) &\geq W(c_0, \int f^*(z_1, x_1)\, \mu(dz_1)) \\ &\geq W(c_0, \int W(c_1, \int f^*(z_2, x_2)\, \mu(dz_2))\, \mu(dz_1)). \end{aligned}$$

Proceed recursively and use Lipschitz continuity of W to prove that

$$f^*(z_0, x_0) \geq (R^N 0)(\mathbf{c}) - \delta^N \int |f^*(z_N, x_N(z^{N-1}))|\, \mu^N(dz^N).$$

Taking the limit as $N \to \infty$ follows $f^*(z_0, x_0) \geq U(\mathbf{c})$ because the second term at the right vanishes by lemma 3(b).

To see that $f^*(z_0, x_0) \leq U(\mathbf{c}) + \varepsilon$ for all $\varepsilon > 0$ for some feasible $\mathbf{c}$ recall that all the steps above can be repeated with equality when we consider a plan generated by the policy correspondence. Indeed, Berge theorem ensures that the policy correspondence defined in (9) (with f^* instead or v) is well defined, compact valued, and upper semicontinuous. Then, theorem 7.6 in Stokey and Lucas (1989) ensures existence of a measurable selection from G. Using this selection it is easy to construct a plan generated by G with $f^*(z_0, x_0) = U(\mathbf{c})$. ■

Observe that f^* has been proven to be the value function $f^* = v$. Proving such statement we have found a plan generated by G with $v(z_0, x_0) = U(\mathbf{c})$. Obviously an optimal plan so that we have:

Corollary 1 *Let (z_0, x_0) be any initial condition and $\mathbf{c} \in \Sigma(z_0, x_0)$ with $\mathbf{x}$ associated. If (8) holds for all z^t and $t \geq 0$ then it is an optimal plan.*

While proposition 3 ensures that v is continuous, Berge theorem ensures that the policy correspondence defined in (9) is well defined, compact valued, and upper semicontinuous. Then two next corollaries follow from corollary 1, the definition of the policy correspondence, and the proof of proposition 3.

Corollary 2 *Every plan generated from G is optimal.*

Corollary 3 *There is an optimal plan.*

One may wonder whether G characterizes every optimal plan. The answer in a stochastic setting is that any optimal plan, if it is not generated by G, it is equal to some generated plan but for zero probability states of nature. The following is theorem 9.4 in Stokey and Lucas (1989) adapted to our case.

Proposition 4 *Let $(\mathbf{c}, \mathbf{x})$ be feasible from $(z_0, x_0) \in Z \times X$. If it is optimal then (8) holds for μ^t almost every z^t and all t.*

Proof: First note that $v(z_0, x_0) = U(\mathbf{c})$. By proposition 2

$$v(z_0, x_0) = W(c_0, \int U(\sigma(z_1, \mathbf{c}))\, \mu(dz_1)).$$

Since $\mathbf{c}$ is optimal, $U(\mathbf{c}) \geq U(\mathbf{c}')$ for all $\mathbf{c}' \in \Sigma(z_0, x_0)$ and therefore

$$W(c_0, \int U(\sigma(z_1, \mathbf{c}))\, \mu(dz_1)) \geq W(c_0', \int U(\sigma(z_1, \mathbf{c}'))\, \mu(dz_1)).$$

Let g be a measurable selection from G and construct a plan $(\mathbf{c}^g, \mathbf{x}^g)$ from this selection with $c_0^g = c_0$ and $x_1^g = x_1$. Two observations: first, since $c_0^g = c_0$ the previous inequality implies that

$$\int U(\sigma(z_1, \mathbf{c}))\, \mu(dz_1)) \geq \int U(\sigma(z_1, \mathbf{c}^g))\, \mu(dz_1);$$

second, for all $z_1 \in Z$ it is true that $\sigma(z_1, \mathbf{c}^g) \in \Sigma(z_1, x_1)$ but since from the first stage on it is a G generated plan corollary 2 ensures that $v(z_1, x_1) = U(\sigma(z_1, \mathbf{c}^g))$ so that

$$U(\sigma(z_1, \mathbf{c})) \leq v(z_1, x_1) = U(\sigma(z_1, \mathbf{c}^g))$$

for all $z_1 \in Z$ and by definition of the value function. These two inequalities together imply that $v(z_1, x_1) = U(\sigma(z_1, \mathbf{c}))$ for μ almost every $z_1 \in Z$ which in turn imply that both integrals with respect to μ are equal. Since $v(z_0, x_0) =$

$W(c_0, \int U(\sigma(z_1, \mathbf{c}))\, \mu(dz_1))$ substitute the integral above by its value to conclude that

$$v(z_0, x_0) = W(c_0, \int v(z_1, x_1)\, \mu(dz_1)).$$

Hence (8) holds for $t = 0$.

To continue note that we have proved that $v(z_1, x_1) = U(\sigma(z_1, \mathbf{c}))$ for μ almost every $z_1 \in Z$ so that the recursive property of U again implies that $v(z_1, x_1) = W(c_1, \int U(\sigma^2(z^2, \mathbf{c}))\, \mu(dz_2)$ for μ almost every $z_1 \in Z$ where σ^2 the composition two times of the shift operator. Following an analogous reasoning to that the paragraph above yields

$$v(z_1, x_1) = W(c_1, \int v(z_2, x_2)\, \mu(dz_2))$$

for μ almost every $z_1 \in Z$. Hence, (8) holds for $t = 1$, and also $v(z_2, x_2) = U(\sigma(z^2, \mathbf{c}))$ for μ^2 almost every $z^2 \in Z^2$ which would be used in the following step. The result follows from proceeding this way recursively. ■

Theorem 1 is proved in propositions 3 and 4 as well as in corollaries 1 and 3. Corollary 2 provides us with a tool to display optimal plans while proposition 4 ensures that this tool fully characterizes optimal plans as non generated ones only differ in null probability states of nature.

References

[1] Aghion, P. and Bolton, P. (1997) "A trickle-down theory of growth and development with debt overhang," Review of Economic Studies, 64, 151-162.

[2] Álvarez, F. and Stokey, N. (1998) "Dynamic programming with homogenous functions," Journal of Economic Theory, 82(1), 167-189.

[3] Becker, R.A. and Boyd III, J.H. (1997) Capital Theory, Equilibrium Analysis and Recursive Utility. Blackwell Publishers.

[4] Bertsekas, D.P. and Shreve, E. (1978) "Mathematical issues in dynamic programming," review paper adapted from "Dynamic programming in Borel spaces," in M. Puterman (ed.) Dynamic Programming and its Applications. Academic Press.

[5] Blackwell, D. (1965) "Discounted dynamic programming," Annals of Mathematical Statistics, 36, 226-235.

[6] Boyd III, J.H. (1990) "Recursive utility and the Ramsey problem," Journal of Economic Theory, 50, 326-345.

[7] Dana, R.A. and Le Van, C. (1990) "Structure of Pareto Optima in an Infinite-Horizon Economy Where Agents Have Recursive Preferences." Journal of Optimization Theory and Applications, 64(2), 269-292.

[8] Díaz-Giménez, J., Quadrini, V. and Ríos-Rull, J.V. (1997) "Dimensions of inequality: Facts on the US distributions of earnings, income, and wealth," Federal Reserve Bank of Minneapolis Quarterly Review, 21(2), 3-21.

[9] Doob, J.L. (1994) Measure Theory. Springer-Verlag.

[10] Durán, J. (2000) "On dynamic programming with unbounded returns," Economic Theory, 15, 339-352.

[11] Dutta, J. and Michel, P. (1998) "The distribution of wealth with imperfect altruism," Journal of Economic Theory, 82, 379-404.

[12] Jones, L.E. and Manuelli, R. (1990) "A convex model of equilibrium growth: theory and policy implications," Journal of Political Economy, 98(5), 1008-1038.

[13] Kreps, D.M. and Porteus, E.L. (1978) "Temporal resolution of uncertainty and dynamic choice theory," Econometrica, 46, 185-200.

[14] Loury, G.C. (1981) "Intergenerational transfers and the distribution of earnings," Econometrica, 49(4), 843-867.

[15] Lucas Jr., R.E. and Stokey, N. (1984) "Optimal growth with many consumers," Journal of Economic Theory, 32, 139-171.

[16] McGrattan, E.R. (1998) "A defense of AK growth models," Federal Reserve Bank of Minneapolis Quarterly Review, 22(4), 13-27.

[17] Ozaki, H. and Streufert, P.A. (1996) "Dynamic programming for non-additive stochastic objectives," Journal of Mathematical Economics, 25, 391-442.

[18] Stokey, N. and Lucas Jr., R.E. with Prescott, E.C. (1989) Recursive Methods in Economic Dynamics. Harvard University Press.

[19] Streufert, P.A. (1996) "Biconvergent stochastic dynamic programming, asymptotic impatience, and average growth," Journal of Economic Dynamics and Control, 20, 385-413.

[20] Wessels, J. (1977) "Markov programming by successive approximations with respect to weighted supremum norms," Journal of Mathematical Analysis and Applications, 58, 326-335.