ISSN 0819-2642
ISBN 0 7340 2553 X

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 897

FEBRUARY 2004

STOCHASTIC GROWTH WITH NONCONVEXITIES: THE OPTIMAL CASE

by

Kazuo Nishimura
&
Ryszard Rudnicki
&
John Stachurski

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

STOCHASTIC GROWTH WITH NONCONVEXITIES: THE OPTIMAL CASE

KAZUO NISHIMURA, RYSZARD RUDNICKI, AND JOHN STACHURSKI

Abstract. This paper studies optimal investment and dynamic behaviour of stochastically growing economies. We assume neither convex technology nor bounded support of the productivity shocks. A number of basic results concerning the investment policy and the Ramsey–Euler equation are established. We also prove a fundamental dichotomy pertaining to optimal growth models perturbed by standard econometric shocks: Either an economy is globally stable or it is globally collapsing to the origin.

1. Introduction

The stochastic optimal growth model (Brock and Mirman, 1972) is a foundation stone of modern marcroeconomic and econometric research. To accommodate the data, however, economists are often forced to go beyond the convex production technology used in these original studies. Nonconvexities lead to technical difficulties which applied researchers would rather not confront. Value functions are in general no longer smooth, optimal policies contain jumps, and the Euler equation may fail. This reality precludes the use of many standard tools. Further, convergence of state variables to a stationary equilibrium is no longer assured. The latter is a starting point of much applied analysis (see, e.g., Kydland and Prescott, 1982; or Long and Plosser, 1983) and fundamental to the rational expectations hypothesis (Lucas, 1986).

This research was partially supported by the Japan Society for the Promotion of Science and the State Committee for Scientific Research (Poland) Grant No. 2 P03A 031 25.

Although nonconvexities are technically challenging, the richer dynamics that they provide help to replicate key time series. For example, nonconvexities often lead to the kind of regime-switching behaviour found in aggregate income data (e.g., Prescott, 2002), or the growth miracles and growth disasters in cross-country income panels. Also, nonconvexities can arise directly from micro-level modeling, taking the form of fixed costs, threshold effects, ecological properties of natural resource systems, economies of scale and scope, network and agglomeration effects, and so on.

The objective of this paper is to investigate in depth the fundamental properties of stochastic nonconvex one-sector models and the series they generate using assumptions which facilitate integration with empirical research (as opposed to analytical convenience).[1]

Previously, in the deterministic case, optimal growth models with nonconvex technology were studied in continuous time by Skiba (1978). In discrete time, Majumdar and Mitra (1982) examined efficiency of intertemporal allocations. Dechert and Nishimura (1983) studied the standard discounted model with convex/concave technology, and characterized the dynamics of the model for every value of the discount factor. More recently, Amir, Mirman and Perkins (1991) used lattice programming techniques to study solutions of the Bellman equation and associated comparative dynamics. Kamihigashi and Roy (2003) study nonconvex optimal growth without differentiability or even continuity.

In the stochastic case, a very rigorous and comprehensive treatment of optimal growth with nonconvex technology is given in Majumdar, Mitra and Nyarko (1989). Amir (1997) studies optimal growth in

[1]We consider only optimal dynamics. There are many studies of nonoptimal competitive dynamics in nonconvex environments. See for example Mirman, Morand and Reffett (2003) and their extensive list of references.

economies that have some degree of convexity. Using martingale arguments, Joshi (1997) analyzes the classical turnpike properties when the production is nonstationary. Schenk-Hoppé (2002) considers dynamic stability of stochastic overlapping generations models with S-shaped production function. Mitra and Roy (2003) study nonconvex renewable resource exploitation and stability of the resource stock.

All of the above papers assume that the shock which perturbs activity in each period has compact support. This assumption makes the analysis more straightforward, but limits applicability to standard econometric models. We assume instead that the distribution of the shock has a density, which may in general have bounded or unbounded support.

The density representation of the shock turns out to be very convenient in proving interiority of the optimal policy and smoothness in the form of Ramsey–Euler equations and related results. Working with these findings and some additional assumptions, we also obtain a fundamental dichotomy for stochastic growth models from this general class. In particular, the economy is either globally stable in a strong sense to be made precise, or globally collapsing to the origin. This result simplifies considerably the range of possible outcomes. We connect the two possibilities to the discount rate, and also provide conditions to determine which outcome prevails for specific parameterizations.

Section 2 introduces the model. Section 3 discusses optimization and properties of the optimal policies. Section 4 considers the dynamics of the processes generated by these policies (i.e., the optimal paths). All of the proofs are given in Section 5 and the appendix.

2. Outline of the Model

Let $\mathbb{R}_+ := [0, \infty)$ and let $\mathscr{B}$ be the Borel subsets of $\mathbb{R}_+$. At the start of each period t a representative agent receives current income $y_t \in$

$\mathbb{R}_+$ and allocates it between current consumption c_t and savings. On current consumption c the agent receives instantaneous utility $u(c)$. For convenience, depreciation is assumed to be total, and current savings determines one-for-one the stock k_t of available capital. Production then takes place, delivering at the start of the next period output

$$y_{t+1} = f(k_t)\varepsilon_t, \tag{1}$$

where ε_t is a shock taking values in $\mathbb{R}_+$.

Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space where uncertainty is generated. In particular, $(\varepsilon_t)_{t=0}^{\infty}$ is a random sequence on this space. The sequence is selected at the start of time according to $\mathbf{P}$ and progressively revealed. When the time t savings decision is made $\varepsilon_0, \ldots, \varepsilon_{t-1}$ are observable.

We assume that each ε_t has the same marginal distribution, which can be represented by density ψ. That is, $\mathbf{P}[\varepsilon_t^{-1}(B)] = \int_B \psi(z)dz$ for all $B \in \mathscr{B}$. Here and in what follows, by density is meant a nonnegative and $\mathscr{B}$-measurable function on $\mathbb{R}_+$ that integrates to unity. For notational convenience the same symbol ψ is used throughout the paper to denote the density function and the distribution of ε, so that $\psi(dz)$ and $\psi(z)dz$ have the same meaning.

The agent seeks to maximize the expectation of a discounted sum of utilities. Future utility is discounted according to $\varrho \in (0, 1)$.

Assumption 2.1. The function u is strictly increasing, twice differentiable on $(0, \infty)$, and satisfies

(U1) $\lim_{c \to 0} u'(c) = \infty$;
(U2) $u''(c) < 0$ for all $c > 0$; and
(U3) u is bounded and $u(0) = 0$.

The Inada condition (U1) is needed to obtain the Ramsey–Euler equation. Strict concavity is critical to the proof of monotonicity of the optimal policy, on which all subsequent results depend. Note that if

u is required to be bounded, then assuming $u(0) = 0$ sacrifices no additional generality.[2]

Assumption 2.2. The production function f is nondecreasing and twice differentiable on $(0, \infty)$. In addition,

(F1) $f(k) = 0$ if and only if $k = 0$;
(F2) $\limsup_{k \to \infty} f'(k) = 0$; and
(F3) $\liminf_{k \to 0} f'(k) > 1$.

Condition (F2) is the usual decreasing returns assumption. Actually for the proofs we require only that f is majorized by an affine function with slope less than one. This is implied by (F2), as can be readily verified from the Fundamental Theorem of Calculus.

Assumption 2.3. The shocks $(\varepsilon_t)_{t \geq 0}$ and their density ψ satisfy

(S1) the sequence $(\varepsilon_t)_{t \geq 0}$ is uncorrelated; and
(S2) $\mathbb{E}^{\mathbf{P}}[\varepsilon_t] = \int z \psi(z) dz = 1$.

Assumption (S2) is just a finite mean assumption—there is no loss of generality in assuming then that the mean is 1.

An economy is defined by the collection (u, f, ψ, ϱ), for which Assumptions 2.1–2.3 are always taken to hold.

By a control policy is meant a function $\sigma \colon \mathbb{R}_+ \ni y \mapsto k \in \mathbb{R}_+$ associating current income to current savings. The policy is said to be feasible if is is $\mathscr{B}$-measurable and $0 \leq \sigma(y) \leq y$ for all y. An initial condition and a feasible policy complete the dynamics of the model (1), determining a stochastic process $(y_t)_{t \geq 0}$ on $(\Omega, \mathscr{F}, \mathbf{P})$, where $y_{t+1} = f(\sigma(y_t))\varepsilon_t$ for all $t \geq 0$.

[2] The theoretical literature uses bounded and unbounded utility functions for dynamic programming. We use the former, because bounded functions are a natural dual pair for probabilities.

Investment behavior is determined by the problem

$$\sup_{\sigma} \mathbb{E}^{\mathbf{P}} \left[\sum_{t=0}^{\infty} \varrho^t u(y_t - \sigma(y_t)) \right], \tag{2}$$

where $\mathbb{E}^{\mathbf{P}}$ denotes integration over Ω with respect to $\mathbf{P}$, an initial condition y_0 is given, and the supremum is over the set of all feasible policies. By (U3) the functional inside the integral is bounded independent of σ, and the supremum always exists. A policy is called optimal if it is feasible and attains the supremum (2).

3. Optimization

In this section we solve the optimization problem by dynamic programming, and characterize the properties of the value function and control policy. To begin, define as usual the value function V by setting $V(y)$ as the real number defined by (2) when $y = y_0$ is the initial condition. Let $b\mathscr{B}$ be the space of real bounded $\mathscr{B}$-measurable functions. Define also the usual operator T mapping $b\mathscr{B}$ into itself by

$$(Tv)(y) = \sup_{0 \leq k \leq y} \left\{ u(y-k) + \varrho \int v[f(k)z]\psi(dz) \right\}. \tag{3}$$

It is well-known that T is a uniform contraction on $b\mathscr{B}$ in the sense of Banach, and that the value function V is the unique fixed point of T in $b\mathscr{B}$.

Lemma 3.1. *For any economy (u, f, ψ, ϱ), the value function V is continuous, bounded and strictly increasing. An optimal policy σ exists. Moreover, if σ is optimal, then*

$$V(y) = u(y - \sigma(y)) + \varrho \int V[f(\sigma(y))z]\psi(dz), \quad \forall y \in \mathbb{R}_+.$$

The proof does not differ from the neoclassical case (see for example Stokey et al., 1989) and is omitted.

As a matter of notation, define

$$\Sigma(y) := \operatorname{argmax}_{0 \leq k \leq y} \left\{ u(y-k) + \varrho \int V[f(k)z]\psi(dz) \right\},$$

so that $y \mapsto \Sigma(y)$ is the optimal correspondence, and σ is an optimal policy if and only if it is a $\mathscr{B}$-measurable selection from Σ.[3]

3.1. **Monotonicity of the policy.** Monotone policy rules play an important role in economics, particularly with regards to the characterization of equilibria. That monotonicity of the optimal investment function holds in one-sector nonconvex growth environments was established by Dechert and Nishimura (1983) and is now well-known. Indeed, monotone controls are a feature of many very general stochastic dynamic environments. See in particular Mirman, Morand and Reffett (2003, Theorem 6 and the discussion in Section 6.2). A simple proof of the following fact is provided for completeness. (Here and below proofs are deferred to Section 5.)

Lemma 3.2. *Let an economy* (u, f, ψ, ϱ) *be given, and let* σ *be a feasible policy. If* σ *is optimal, then it is nondecreasing on* $\mathbb{R}_+$*.*

Put differently, one cannot construct a measurable selection from the optimal correspondence Σ that is not nondecreasing. (In contrast to the neoclassical case, in nonconvex models consumption is not generally monotone with income.) It should be emphasized that Lemma 3.2 holds under *much weaker* conditions—in particular without interiority-type assumptions (Mirman, Morand and Reffett, 2003), continuity (Kamihigashi and Roy, 2003) and so on.

One supposes that as ϱ decreases—increasing the rate at which the future is discounted—the propensity to save will fall. The following result was established for the stochastic neoclassical case in Danthine

[3] Regarding the existence of a measurable selection σ, see, for example, Hopenhayn and Prescott (1992, Lemma 2).

and Donaldson (1981, Theorem 5.1), and in the nonconvex, deterministic case by Amir, Mirman and Perkins (1991). The second paper gives an attractive proof using lattice programming. Here we provide a very elementary proof.

Lemma 3.3. *The optimal policy is nondecreasing in the discount factor* ϱ*, in the sense that if* (u, f, ψ, ϱ_0) *and* (u, f, ψ, ϱ_1) *are two economies, and if* σ_0 *(resp.* σ_1*) is optimal for the former (resp. latter), then* $\varrho_1 \geq \varrho_0$ *implies* $\sigma_1 \geq \sigma_0$ *pointwise on* $\mathbb{R}_+$*.*

Moreover, any sequence of optimal policies not only decreases, but also converges to zero as ϱ does—in fact uniformly on compacts:

Lemma 3.4. *For* u*,* f *and* ψ *given, let* (ϱ_n) *be a sequence of discount factors in* $(0, 1)$*, and for each* n *let* σ_n *be a corresponding optimal policy. If* $\varrho_n \downarrow 0$*, then* $\sigma_n \downarrow 0$ *pointwise, and the convergence is uniform on compact sets.*

3.2. **Derivative characterization of the policy.** Optimal behavior in growth models is usually characterized by the Ramsey–Euler equation—an intuitive and tractable intertemporal arbitrage condition. In stochastic models, where sequential arguments are unavailable, the obvious path to the Ramsey–Euler equation is via differentiability of the value function and a well-known envelope condition (Mirman and Zilcha, 1975, Lemma 1). In the case of the one-sector neoclassical model, all of these results were already established and carefully investigated by Mirman and Zilcha (1975) and others.

Further progress was made by Blume, Easley and O'Hara (1982), who demonstrated differentiability of the optimal policy under convexity and absolute continuity of the shock by way of the implicit function theorem. Amir (1997) extended these results to a weaker convexity requirement.

Without any convexity, however, there may be jumps in the optimal policy, which in turn affect the smoothness of the value function. The validity of the Ramsey–Euler characterization is by no means clear. However, Dechert and Nishimura (1983, Theorem 6, Lemma 8) showed that in their model the value function has both left and right derivatives at every point, and that these agree off an at most countable set. The intuition is that nondifferentiability of the value function coincides pointwise with jumps in optimal investment. But by Lemma 3.2, the only optimal jumps are increases. To each jump, then, can be associated a unique rational, which precludes uncountability.

These results were extended to the stochastic case by Majumdar, Mitra and Nyarko (1989). In addition to the above results concerning the value function, they were able to show for the first time that the Ramsey–Euler equations holds *everywhere*, irrespective of jumps in the optimal policy.

Although their findings remain an important, they require that shocks are supported on a compact interval bounded away from zero, which excludes standard empirical formulations. Also, they assume the existence of a neighborhood of zero in which output strictly exceeds capital input with probability one. In the present paper a different approach is used, starting from the essential idea of Blume, Easley and O'Hara (1982), but without convexity or compact state. From this we prove interiority of the policy and the Ramsey–Euler equation for standard econometric shocks.

Assumption 3.1. The shock ε_t is such that

(S3) the density ψ is continuously differentiable on $(0, \infty)$, and

(S4) the integral $\int z|\psi'(z)|dz$ is finite.

The set of densities satisfying (S3) and (S4) is norm-dense in the set of all densities when the later are considered as a subset of $L_1(\mathbb{R}_+)$.

They also hold for many standard econometric shocks on $\mathbb{R}_+$, such as the lognormal distribution. We these assumptions in hand we can establish the following without convexity or bounded shocks.

Proposition 3.1. *Let (u, f, ψ, ϱ) satisfy Assumptions 2.1–3.1.*

1. *If policy σ is optimal, then it is interior. That is, $0 < \sigma(y) < y$ for all $y \in (0, \infty)$.*
2. *The value function V has right and left derivatives V'_- and V'_+ everywhere on $(0, \infty)$.*
3. *If policy σ is optimal, then it satisfies $V'_-(y) \leq u'(y - \sigma(y)) \leq V'_+(y)$ for all $y \in (0, \infty)$.*
4. *The functions V'_- and V'_+ disagree on an at most countable subset of $\mathbb{R}_+$.*

In the stochastic nonconvex case, Part 1 of Proposition 3.1 was proved by Majumdar, Mitra and Nyarko (1989, Theorem 4). Their proof requires that the shock has compact support bounded away from zero, and there exists an $a > 0$ such that $f(k)\varepsilon > k$ with probability one whenever $k \in (0, a)$. Part 2 was proved in the deterministic case by Dechert and Nishimura, as was Part 4 (1983, Theorem 6 and Lemma 8).[4] Part 3 is due in the stochastic neoclassical case to Mirman and Zilcha (1975, Lemma 1), and the proof for the nonconvex case is the same.[5]

Corollary 3.1. *For a given economy (u, f, ψ, ϱ), any two optimal policies are equal almost everywhere.*

[4]On Part 2 see also Askri and Le Van (1998, Proposition 3.2) and Mirman, Morand and Reffet (2003).

[5]Differentiability of the value function for the stochastic neoclassical growth model was first established by Mirman and Zilcha (1975, Lemma 1). They argued that if V is concave on some open interval, then the subdifferentials exist everywhere on that interval, and $V'_+ \leq V'_-$. If follows from Part 3 of the Proposition, then, that concavity immediately gives differentiability, and, moreover, $V'(y) = u'(y - \sigma(y))$.

Proof. Immediate from Parts 3 and 4. □

It will turn out that under the maintained assumptions, differences on null sets do not really concern us (see Lemma 4.1). So we can in some sense talk about *the* optimal policy (i.e., when a.e.-equivalent policies are identified).

It turns out that even with the maintained assumptions the Ramsey–Euler equation continues to hold.

Proposition 3.2. *Let Assumptions 2.1–3.1 hold. If σ is optimal for (u, f, ψ, ϱ), then for all $y > 0$,*

$$u'(y - \sigma(y)) = \varrho \int u'[f(\sigma(y))z - \sigma(f(\sigma(y))z)]f'(\sigma(y))z\psi(z)dz.$$

Using Proposition 3.2 we can strengthen the monotonicity result for the optimal policy (Lemma 3.2). The proof is straightforward and is omitted.[6]

Corollary 3.2. *For a given economy (u, f, ψ, ϱ), every optimal policy is strictly increasing.*

With these restrictions it becomes possible to investigate in detail the dynamical behavior of the optimal paths.

4. Dynamics

Next we discuss the dynamics of the stochastic process $(y_t)_{t \geq 0}$. For the nonconvex deterministic case a detailed characterization of dynamics was given by Dechert and Nishimura (1983). Not surprisingly, for some parameter values multiple equilibria obtain. For the convex stochastic growth model, Mirman (1970) and Brock and Mirman (1972) proved that the sequence of marginal distributions for the process converge to

[6] Strict concavity of u is necessary here. See for example Mirman, Morand and Reffett (2003, Section 6.2).

a unique limit independent of the initial condition. Subsequently this problem has been treated by many authors.[7]

For the stochastic convex model the stability proofs require convex technology and infinite marginal product of capital at the origin. However, not all environments are convex, and there is little empirical evidence to suggest that in the aggregate production function an infinite marginal productivity of capital at the origin is certain, or even likely. Indeed, casual observation shows that not all economies converge upwards on a stable growth path. Stagnation and collapse also occur.

When the Brock-Mirman conditions are weakened the potential for instability arises. Which kind of dynamical behavior prevails will be determined by a complex interaction between preferences, technology and the investment behavior of agents. Regarding asymptotic stability of optimal stochastic growth models without convexity, relatively little is known.[8] Kamihigashi (2003) shows that even for shocks which are only stationary and ergodic, sufficiently adverse distributions lead to a.s. convergence to zero for every feasible policy. Mitra and Roy (2003) give interesting discussions of extinction and conservation in renewable resource models. Joshi (1997) uses monotonicity and martingale arguments to prove various turnpike results.

Below it is shown that optimal processes satisfy a fundamental dichotomy. Either they are globally stable or globally collapsing to the origin, independent of the initial condition. This result reduces considerably the possible range of asymptotic outcomes. For example, path depedence never holds. More importantly, global stability can now be established by showing only that an economy does not collapse to the origin. The proof is based on the Foguel Alternative for Markov

[7]See for example Stachurski (2002) and references.

[8]This is mainly because the properties of the optimal policies are difficult to determine, rather than any inherent difficultly in analyzing nonlinear or discontinuous stochastic dynamics.

chains (Foguel 1969, Rudnicki 1995). Some conditions are provided to distinguish between the two possibilities.

To begin, let $\mathscr{P}$ be the set of probability measures on $(\mathbb{R}_+, \mathscr{B})$. For a fixed policy σ and initial condition y_0, we consider the evolution of the income process $(y_t)_{t\geq 0}$ satisfying $y_{t+1} = f(\sigma(y_t))\varepsilon_t$, and the corresponding sequence of marginal distributions $(\varphi_t)_{t\geq 0} \subset \mathscr{P}$.[9] By (S1) the process is Markovian, with y_t independent of ε_t. In particular, for any bounded Borel function $h\colon \mathbb{R}_+ \to \mathbb{R}$,

$$\mathbb{E}^{\mathbf{P}} h(y_{t+1}) = \mathbb{E}^{\mathbf{P}} h[S(y_t)\varepsilon_t] = \int\int h[S(y)z]\psi(dz)\varphi_t(dy).$$

Specializing to the case $h = \mathbf{1}_B$ and using $y_t \sim \varphi_t$ gives the recursion

$$\varphi_{t+1}(B) = \int \left[\int \mathbf{1}_B[S(y)z]\psi(z)dz\right] \varphi_t(dy). \tag{4}$$

When σ is optimal, the sequence of marginal distributions (φ_t) defined inductively by (4) is called an optimal path. Evidently it depends on σ and the initial condition $y_0 \sim \varphi_0$, which is taken as data.

If $y_0 = 0$ the dynamics require no additional investigation. Henceforth by an *initial condition* is meant a random variable y_0 such that $\mathbf{P}\{y_0 > 0\} = 1$ holds. This convention makes the results a bit neater, and is maintained throughout the proofs without further comment.

When studying convergence of probabilities two topologies are commonly used. One is the so-called weak topology, under which distribution functions converge if and only if they converge pointwise at all continuity points.[10] The other is the norm topology, or strong topology, generated by the total variation norm. Under the latter, the distance

[9]As before, $(y_t)_{t\geq 0}$ is a stochastic process on $(\Omega, \mathscr{F}, \mathbf{P})$. By the marginal distribution $\varphi_t \in \mathscr{P}$ of y_t is meant its distribution on $\mathbb{R}_+$ in the usual sense. Precisely, $\varphi_t := \mathbf{P} \circ y_t^{-1}$, the image measure induced on $(\mathbb{R}_+, \mathscr{B})$ by y_t.

[10]It is the smallest topology on $\mathscr{P}$ making the functionals $\mathscr{P} \ni \mu \mapsto \int g d\mu \in \mathbb{R}$ continuous for each $g \in C_b(\mathbb{R}_+)$. Here $C_b(\mathbb{R}_+)$ is the continuous bounded functions on $\mathbb{R}_+$.

between μ and ν in $\mathscr{P}$ is the supremum of $\sum_n^N |\mu(B_n) - \nu(B_n)|$ over all finite measurable partitions of $\mathbb{R}_+$.

Definition 4.1. Let an economy (u, f, ψ, ϱ) be given, and let σ be an optimal policy. Following Mirman (1970), an equilibrium or stochastic steady state for (u, f, ψ, ϱ) is a measure $\varphi^* \in \mathscr{P}$, such that $\varphi^*(\{0\}) = 0$ and

$$\int \left[\int \mathbf{1}_B[f(\sigma(y))z]\psi(dz) \right] \varphi^*(dy) = \varphi^*(B), \quad \forall B \in \mathscr{B}, \tag{5}$$

where $\mathbf{1}_B$ is the indicator function of B. The policy σ is called globally stable if for σ there is a unique equilibrium φ^*, and the optimal path (φ_t) generated by σ and starting at φ_0 satisfies $\varphi_t \to \varphi^*$ in the norm topology as $t \to \infty$ for all initial conditions φ_0. The economy (u, f, ψ, ϱ) is called globally stable if every optimal policy is globally stable.

Equation (5) should be understood as follows. The integral term inside the brackets is the probability that income is in B next period, given that it is currently equal to y. The outer integral averages this over all y, weighted by φ^*. Thus the left hand side is the probability that income is in B next period given that it is currently distributed according to φ^*. If this is again equal to $\varphi^*(B)$ then the economy is in equilibrium.

The stability condition defined above is a particularly strong one. It implies many standard stability conditions for Markov processes, such as recurrence, and also convergence of the marginal distributions in the weak topology.[11]

[11] In the present case it also implies uniform convergence of distribution functions, which is the criterion of Brock and Mirman (1972). See Dudley (2002, p. 389).

For stochastic growth instability has been studied less than stability. There are various notions which capture instability; we borrow a relatively strong one from the Markov process literature referred to as sweeping.[12]

Definition 4.2. Let an economy (u, f, ψ, ϱ) be given, and let σ be an optimal policy. Let $\mathscr{B}_0 \subset \mathscr{B}$. In general, the Markov process generated by the policy σ is called sweeping with respect to $\mathscr{B}_0$ if each optimal path (φ_t) generated by σ satisfies $\varphi_t(A) \to 0$ as $t \to \infty$ for every $A \in \mathscr{B}_0$ and every initial condition φ_0. Here we say that policy σ is globally collapsing to the origin if it is sweeping with respect the the collection of intervals $[a, \infty)$, $a > 0$. Also, (u, f, ψ, ϱ) will be called globally collapsing to the origin if every optimal policy is.

The following result indicates that there is a fundamental dichotomy for the dynamic behavior of the economy. In the proofs monotonicity and interiority of the optimal policy play key roles.

Assumption 4.1. Density ψ is strictly positive (Lebesgue almost) everywhere on $\mathbb{R}_+$

Proposition 4.1. *Let an economy (u, f, ψ, ϱ) be given. If in addition to Assumptions 2.1–3.1, Assumption 4.1 also holds, then there are only two possibilities. Either*

1. (u, f, ψ, ϱ) *is globally stable, or*
2. (u, f, ψ, ϱ) *is globally collapsing to the origin.*

Remark. Assumption 4.1 can be weakened significantly (Rudnicki, 1995, Lemma 3 and Theorem 2), but it holds for many standard econometric shocks so we maintain it.

Thus for stochastic optimal growth models with these assumptions multiple equilibria are never observed, regardless of nonconvexities in

[12]See, for example, Lasota and Mackey (1994, Section 5.9).

production technology. Instead long run outcomes are completely determined by the structure of the model, and historical conditions are asymptotically irrelevant. However, the equilibrium distribution may well be multi-modal, concentrated on areas that are locally attracting on average.

Nonconvex technology introduces the possibility that many optimal policies exist for the one economy. For the deterministic nonconvex case it has been shown (Dechert and Nishimura, 1983, Lemma 6) that different optimal trajectories can have very different dynamics, even from the same initial condition. For our stochastic model this is not possible:

Lemma 4.1. *Let an economy (u, f, ψ, ϱ) be given. If one optimal policy (i.e., measurable selection from Σ) is globally asymptotically stable, then every optimal policy (selection from Σ) is, and hence so is (u, f, ψ, ϱ). Conversely, if one optimal policy is globally collapsing to the origin, then every optimal policy is, and hence so is (u, f, ψ, ϱ).*

We have seen that an increase in the discount rate (a decrease in ϱ) is associated with less savings and investment, which in turn should increase the likelihood of collapse to the origin. Conversely, lower discount rates (higher ϱ) should increase the likelihood that the economy is stable. Precisely,

Lemma 4.2. *For economies $E_0 := (u, f, \psi, \varrho_0)$ and $E_1 := (u, f, \psi, \varrho_1)$ with $\varrho_0 \leq \varrho_1$, the following implications hold.*

1. *If E_1 is globally collapsing to the origin, then so is E_0.*
2. *If E_0 is globally asymptotically stable, then so is E_1.*

Combining the above results we can deduce that the dynamic behavior of the stochastic optimal growth model has only three possible types. Precisely,

Proposition 4.2. *For u, f, and ψ given, either*

1. *(u, f, ψ, ϱ) is globally stable for all $\varrho \in (0, 1)$,*
2. *(u, f, ψ, ϱ) is globally collapsing for all $\varrho \in (0, 1)$, or*
3. *there is a $\hat{\varrho} \in (0, 1)$ such that (u, f, ψ, ϱ) is globally stable for all $\varrho > \hat{\varrho}$, and globally collapsing for all $\varrho < \hat{\varrho}$.*

We emphasize that under the current hypotheses one cannot rule out the possibility that the economy is globally stable or globally collapsing for every $\varrho \in (0, 1)$. For example, Kamihigashi (2003) shows that very general one-sector growth models converge almost surely to zero when $f'(0) < \infty$ and shocks are sufficiently volatile. Determining which of the above three possibilities holds, then, requires careful study of individual models. To this end we prove the following dynamical systems result which gives sufficient conditions against which different model primitives can be tested. It shows that the difference between global stability and global instability depends only on the behavior of the model in the neighborhood of the origin.

Assumption 4.2. Density ψ satisfies $\mathbb{E}|\ln \varepsilon| = \int |\ln z| \psi(dz) < \infty$.

Proposition 4.3. *Let an economy (u, f, ψ, ϱ) be given, and let σ be an optimal policy. Suppose that Assumptions 2.1–4.2 hold. Define*

$$p := \limsup_{y \to 0} \frac{f(\sigma(y))}{y}, \quad q := \liminf_{y \to 0} \frac{f(\sigma(y))}{y}.$$

1. *If $p < \exp(\mathbb{E} \ln \varepsilon)$, then (u, f, ψ, ϱ) is globally collapsing to the origin.*
2. *If $q > \exp(\mathbb{E} \ln \varepsilon)$, then (u, f, ψ, ϱ) is globally stable.*

Also, in the light of Lemma 3.4, one might suspect that even in the situation where an economy is globally stable for every ϱ, the stationary distribution will become more and more concentrated around the origin when the discount rate becomes very large ($\varrho \downarrow 0$). In this connection,

Proposition 4.4. *Let u, f and ψ be given. Suppose that (u, f, ψ, ϱ) is globally stable for all $\varrho \in (0, 1)$. If $\varrho_n \to 0$, then $\varphi_n^* \to \delta_0$ in the weak topology, where φ_n^* is the stationary distribution corresponding to ϱ_n, and δ_0 is the probability measure concentrated at zero.*

Remark. Norm (as opposed to weak) convergence is impossible here, because— as is clear from the proofs—the stationary distribution must be a density, in which case δ_0 and φ_n^* are mutually singular, and $\|\delta_0 - \varphi_n^*\| = 2$ for all n.

5. Proofs

In the proofs, $L_1(X)$ refers as usual to all integrable Borel functions on given space X, and $C^n(X)$ is the n times continuously differentiable functions.

5.1. **Monotonicity.**
The proof of monotonicity of the optimal policy is as follows.

Proof of Lemma 3.2. Let σ be optimal, and take any nonnegative $y \leq y'$. If $y = y'$ then monotonicity is trivial. Suppose the inequality is strict. By way of contradiction, suppose that $\sigma(y) > \sigma(y')$. Define $c := y - \sigma(y)$, $c' := y' - \sigma(y')$, and $\hat{c} := \sigma(y) - \sigma(y') > 0$. Note first that

$$c' - \hat{c} = y' - \sigma(y) > y - \sigma(y) = c \geq 0. \tag{6}$$

Also, since $c + \hat{c} + \sigma(y') = y$, we have

$$u(c) + \varrho \int V[f(\sigma(y))z]\psi(dz) \geq u(c + \hat{c}) + \varrho \int V[f(\sigma(y'))z]\psi(dz),$$

and since $c' - \hat{c} + \sigma(y) = y'$,

$$u(c') + \varrho \int V[f(\sigma(y'))z]\psi(dz) \geq u(c' - \hat{c}) + \varrho \int V[f(\sigma(y))z]\psi(dz).$$

$$\therefore \quad u(c') - u(c' - \hat{c}) \geq u(c + \hat{c}) - u(c).$$

As $c' - \hat{c} > c$ by (6), this contradicts the strict concavity of u. □

Proof of Lemma 3.3. Pick any $y \geq 0$. Let $k_0 := \sigma_0(y)$ and $k_1 := \sigma_1(y)$. By definition,

$$u(y-k_0) + \varrho_0 \int V(f(k_0)z)\psi(dz) \geq u(y-k_1) + \varrho_0 \int V(f(k_1)z)\psi(dz)$$

and

$$u(y-k_1) + \varrho_1 \int V(f(k_1)z)\psi(dz) \geq u(y-k_0) + \varrho_1 \int V(f(k_0)z)\psi(dz).$$

Multiplying the first inequality by ϱ_1 and the second by ϱ_0 and adding gives

$$\varrho_1 u(y-k_0) + \varrho_0 u(y-k_1) \geq \varrho_1 u(y-k_1) + \varrho_0 u(y-k_0).$$

$$\therefore \quad (\varrho_1 - \varrho_0)(u(y-k_0) - u(y-k_1)) \geq 0.$$

$$\therefore \quad \varrho_1 \geq \varrho_2 \implies u(y-k_0) - u(y-k_1) \geq 0 \implies k_1 \geq k_0.$$

□

Proof of Lemma 3.4. Since u is concave, for any $y > 0$ and any $k \leq y$,

$$u(y-k) \leq u(y) - u'(y)k. \tag{7}$$

Also, since $u(y) \leq M < \infty$ for all y,

$$V(y) := \sup_{\sigma} \mathbb{E}^{\mathbf{P}} \left[\sum_{t=0}^{\infty} \varrho^t u(y_t - \sigma(y_t)) \right] \leq \frac{1}{1-\varrho} M. \tag{8}$$

Since $\sigma(y) = 0$ is feasible,

$$u(y - \sigma(y)) + \varrho \int V(f(\sigma(y))z)\psi(dz) \geq u(y) + \varrho \int V(f(0)z)\psi(dz) = u(y).$$

$$\therefore \quad u(y) - u(y-\sigma(y)) \leq \varrho \int V(f(\sigma(y))z)\psi(dz) \leq \frac{\varrho}{1-\varrho} M.$$

Using the bound (7) gives us

$$u'(y)\sigma(y) \leq \frac{\varrho}{1-\varrho} M, \quad \forall y > 0.$$

$$\therefore \quad \sigma(y) \leq \frac{\varrho}{1-\varrho} \frac{M}{u'(y)} := b(y;\varrho).$$

The function $y \to b(y,\varrho)$ is continuous and converges pointwise to zero as $\varrho \to 0$. The statement follows (uniform convergence on compact sets is by Dini's Theorem). □

5.2. **The Ramsey–Euler equation.** Next Propositions 3.1 and 3.2 are established. We use the following lemma, which can by thought of as a kind of convolution argument designed to verify precisely the conditions necessary for the Ramsey–Euler equation to hold. The proof is rather long, and is relegated to the appendix.

Lemma 5.1. *Let g and h be nonnegative real functions on $\mathbb{R}$. Define*

$$\mu(r) := \int_{-\infty}^{\infty} h(x+r)g(x)\,dx. \tag{9}$$

Consider the following conditions:

(i) $g \in L_1(\mathbb{R}) \cap C^1(\mathbb{R})$, $g' \in L_1(\mathbb{R})$
(ii) *h is bounded*
(iii) *h is nondecreasing*
(iv) *h is absolutely continuous on closed intervals*
(v) *h' is bounded on compact subsets of $\mathbb{R}$,*

where h' is defined as the derivative of h when it exists and zero elsewhere. If (i) *and* (ii) *hold, then $\mu \in C^1(\mathbb{R})$, and*

$$\mu'(r) = -\int_{-\infty}^{\infty} h(x+r)g'(x)\,dx. \tag{10}$$

If, in addition, (iii)–(v) *hold, then μ' also has the representation*

$$\mu'(r) = \int_{-\infty}^{\infty} h'(x+r)g(x)\,dx. \tag{11}$$

Remark. Note that higher order derivatives are immediate if g has high order derivatives that are all integrable. In the first part of the proof, where differentiability and the representation $\mu'(r) = -\int h(x+r)g'(x)dx$ are established we do not use nonnegativity of g—it may be any real function. So now suppose that g is twice differentiable, and that $g'' \in L_1(\mathbb{R})$. Then by applying the same result, this time using g' for g, differentiability of μ' is verified.

To prove Proposition 3.1, the following preliminary observation is important. (Assume the hypotheses of that proposition.)

Lemma 5.2. *If V is the value function for (u, f, ψ, ϱ), then $k \mapsto \int V[f(k)z]\psi(z)dz$ is continuously differentiable on the interior of $\mathbb{R}_+$.*

Proof. By a simple change of variable,

$$\int_0^\infty V[f(k)z]\psi(z)dz = \int_{-\infty}^\infty V[\exp(\ln f(k) + x)]\psi(e^x)e^x\, dx.$$

Let $h(x) := V[\exp(x)]$, $g(x) := \psi(e^x)e^x$, and let μ be defined as in (9). Then $\int V[f(k)z]\psi(z)dz = \mu[\ln f(k)]$. Regarding μ, conditions (i) and (ii) of Lemma 5.1 are satisfied by (U3) and (S2), (S4). Hence $\int V[f(k)z]\psi(z)dz$ is continuously differentiable as claimed. □

Now let us consider the interiority result.

Proof of Proposition 3.1, Part 1. Pick any $y > 0$. Consider first the claim that $\sigma(y) \neq 0$. Suppose instead that $0 \in \Sigma(y)$, so that

$$V(y) = u(y) - \varrho \int V[f(0)z]\psi(dz) = u(y), \tag{12}$$

where we have used $u(0) = 0$ in (U3). Define also

$$V_\xi := u(y - \xi) + \varrho \int V[f(\xi)z]\psi(dz), \tag{13}$$

where ξ is a positive number less than y. By (F3), there exists a $\delta > 0$ such that $f(\xi) > \xi$ whenever $\xi < \delta$. Therefore,

$$V_\xi \geq u(y - \xi) + \varrho \int V(\xi z)\psi(dz), \quad \forall \xi < \delta. \tag{14}$$

In addition, $V \geq u$ everywhere on $\mathbb{R}_+$. Using this bound along with (12) and (14) gives

$$0 \leq \frac{V(y) - V_\xi}{\xi} \leq \frac{u(y) - u(y-\xi)}{\xi} - \varrho \int \frac{u(\xi z)}{\xi}\psi(dz), \quad \forall \xi < \delta. \tag{15}$$

Take a sequence $\xi_n \downarrow 0$. If $H_n(z) = u(\xi_n z)/\xi_n$, then $H_n \geq 0$ on $\mathbb{R}_+$ and $H_{n+1}(z) \geq H_n(z)$ for all z and all n. Moreover $\lim_{n\to\infty} H_n = \infty$ almost everywhere. By the Monotone Convergence Theorem, then,

$$\lim_{n\to\infty} \int \frac{u(\xi_n z)}{\xi_n}\psi(dz) = \int \infty\, \psi(dz) = \infty,$$

which induces a contradiction in (15).

Now consider the claim that $\sigma(y) \neq y$. Let

$$v(k) := u(y-k) + w(k), \quad w(k) := \varrho \int V[f(k)z]\psi(dz), \quad k \in [0, y].$$

If $y \in \Sigma(y)$, then for all positive ε,

$$0 \leq \frac{v(y) - v(y-\varepsilon)}{\varepsilon} = -\frac{u(\varepsilon)}{\varepsilon} + \frac{w(y) - w(y-\varepsilon)}{\varepsilon}. \tag{16}$$

Since $w(k)$ is differentiable at y (Lemma 5.2), the second term on the right-hand side converges to a finite number as $\varepsilon \downarrow 0$. In this case clearly there will be a contradiction of inequality (16). This completes the proof that $y \notin \Sigma(y)$. □

Proof of Proposition 3.1, Part 2. Regarding the existence of left and right derivatives, pick any $y > 0$, any $\xi_n \downarrow 0$, $\xi_n > 0$, and any optimal policy σ. By monotonicity, $\sigma(y + \xi_n)$ converges to some limit k_+, and the value k_+ is independent of the choice of sequence (ξ_n). Moreover, upper hemi-continuity of Σ implies that k_+ is maximal at y. It follows from this and interiority of optimal policies that $0 < k_+ < y$ and

$$V(y) = u(y - k_+) + \varrho \int V[f(k_+)z]\psi(dz).$$

Also, for all $n \in \mathbb{N}$,

$$\begin{aligned} V(y + \xi_n) &= u(y + \xi_n - \sigma(y + \xi_n)) + \varrho \int V[f(\sigma(y+\xi_n))z]\psi(dz) \\ &\geq u(y - k_+ + \xi_n) + \varrho \int V[f(k_+)z]\psi(dz). \end{aligned}$$

$$\therefore \quad u(y - k_+ + \xi_n) - u(y - k_+) \leq V(y + \xi_n) - V(y), \quad \forall n \in \mathbb{N}.$$

On the other hand, since $\sigma(y + \xi_n) \downarrow k_+ < y$, there exists an $N \in \mathbb{N}$ such that

$$V(y) \geq u(y - \sigma(y + \xi_n)) + \varrho \int V[f(\sigma(y+\xi_n))z]\psi(dz), \quad \forall n \geq N.$$

$$\therefore \quad V(y+\xi_n) - V(y) \leq u(y + \xi_n - \sigma(y+\xi_n)) - u(y - \sigma(y+\xi_n)), \quad \forall n \geq N.$$

$$\therefore \quad V(y+\xi_n) - V(y) \leq u'(y - \sigma(y+\xi_n))\xi_n, \quad \forall n \geq N,$$

where the last inequality is by concavity of u. In summary, then,

$$u(y - k_+ + \xi_n) - u(y - k_+) \leq V(y+\xi_n) - V(y) \leq u'(y - \sigma(y+\xi_n))\xi_n$$

for all n sufficiently large. Dividing through by $\xi_n > 0$ and taking limits gives $V'_+(y) = u'(y - k_+)$, which is of course finite by $k_+ < y$.[13]

[13] We are using continuity of u', which is guaranteed by twice differentiability.

Now consider the analogous argument for V'_-. Let y, (ξ_n) and σ be as above. Again, as σ is monotone, $\sigma(y - \xi_n) \uparrow k_-$, where k_- is independent of the precise sequence (ξ_n), maximal at y and satisfies $0 < k_- < y$. Since $k_- > 0$, then sequence $\sigma(y - \xi_n)$ will be positive for large enough n and we can assume this is so for all n. By maximality,

$$V(y) = u(y - k_-) + \varrho \int V[f(k_-)z]\psi(dz).$$

Also, since $k_- < y$, there exists an $N \in \mathbb{N}$ with $k_- \leq y - \xi_n$ for all $n \geq N$. Hence, $\forall n \geq N$,

$$\begin{aligned} V(y - \xi_n) &= u(y - \xi_n - \sigma(y - \xi_n)) + \varrho \int V[f(\sigma(y - \xi_n))z]\psi(dz) \\ &\geq u(y - k_- - \xi_n) + \varrho \int V[f(k_-)z]\psi(dz). \end{aligned}$$

$$\therefore \quad u(y - k_- - \xi_n) - u(y - k_-) \leq V(y - \xi_n) - V(y), \quad \forall n \geq N.$$

One the other hand, since $0 < \sigma(y - \xi_n) \uparrow k_- < y$,

$$V(y) \geq u(y - \sigma(y - \xi_n)) + \varrho \int V[f(\sigma(y - \xi_n))z]\psi(dz), \quad \forall n \in \mathbb{N}.$$

$$\therefore \quad V(y - \xi_n) - V(y) \leq u(y - \xi_n - \sigma(y - \xi_n)) - u(y - \sigma(y - \xi_n)), \quad \forall n \in \mathbb{N}.$$

$$\therefore \quad V(y - \xi_n) - V(y) \leq -u'(y - \sigma(y - \xi_n))\xi_n, \quad \forall n \in \mathbb{N},$$

where again the last inequality is by concavity of u. Putting the inequalities together gives

$$u(y - k_- - \xi_n) - u(y - k_-) \leq V(y - \xi_n) - V(y) \leq u'(y - \sigma(y - \xi_n))(-\xi_n)$$

for all n sufficiently large. Dividing through by $-\xi_n$ and taking limits gives $V'_-(y) = u'(y - k_-)$. □

Proof of Proposition 3.1, Part 3. The proof is identical to that given in Mirman and Zilcha (1975, Lemma 1). □

Proof of Proposition 3.1, Part 4. The proof is essentially the same as that Majumdar, Mitra and Nyarko (1989, Lemma 4). Briefly, it is clear from the proof of Part 2 of Proposition 3.1 that $V'_-(y)$ and $V'_+(y)$ will agree whenever $\Sigma(y)$ is a singleton. If y_1 and y_2 are any two distinct points where Σ is multi-valued, then $\Sigma(y_1)$ and $\Sigma(y_2)$ can intersect at at most one point, otherwise we can construct a non-monotone optimal policy, contradicting Lemma 3.2.

It follows that for each y where $\Sigma(y)$ is multi-valued, $\Sigma(y)$ can be allocated a unique rational number. $\square$

Next we come to the proof of the Ramsey–Euler equation. We need the following lemma, which was first proved (under different assumptions) by Majumdar, Mitra and Nyarko (1989, Lemma 2A).

Lemma 5.3. *For every compact* $K \subset (0, \infty)$, $\inf\{y - \sigma(y) : y \in K\}$ *is strictly positive.*

Proof. Suppose to the contrary that on some compact set $K \subset (0, \infty)$, there exists for each n a y_n with $\sigma(y_n) > y_n - 1/n$. By compactness (y_n) has a convergent subsequence, and without loss of generality we assume that the whole sequence converges to $y^* \in K$. The bounded sequence $\sigma(y_n)$ itself has a convergent subsequence $\sigma(y_{n(i)}) \to k^*$ as $i \to \infty$. Since the subsequence $(y_{n(i)})$ converges to y^* too, k^* is optimal at y^* by upper hemicontinuity. But then $y^* - \frac{1}{n(i)} \leq k^* \leq y^*$ for all $i \in \mathbb{N}$. This contradicts the interiority of the optimal policy, which has already been established. $\square$

The next lemma is fundamental to our results.

Lemma 5.4. *Define* V' *to be the derivative of* V *when it exists and zero elsewhere. For all* $k > 0$,

$$\frac{d}{dk} \int V[f(k)z]\psi(z)dz = \int V'(f(k)z)f'(k)z\psi(z)dz.$$

Proof. We change variables to shift the problem to the real line. Our objective is to apply Lemma 5.1. Let

$$w(k) := \int V[f(k)z]\psi(z)dz.$$

As before, we can use a change of variable to obtain

$$w(k) = \int_{-\infty}^{\infty} V(f(k)e^x)\psi(e^x)e^x dx = \int_{-\infty}^{\infty} h(x + \ln f(k))g(x)dx,$$

where $g(x) := \psi(e^x)e^x$ and $h(x) := V(e^x)$. All of the hypotheses of Lemma 5.1 are satisfied.[14] Therefore, using the representation (11),

$$\begin{aligned} w'(k) &= \frac{f'(k)}{f(k)} \int_{-\infty}^{\infty} h'(x + \ln f(k)) g(x) dx \\ &= f'(k) \int_{-\infty}^{\infty} V'(e^x f(k)) e^x g(x) dx. \end{aligned}$$

Changing variables again gives the desired result:

$$w'(k) = \int_0^{\infty} V'(f(k)z) f'(k) g(\ln z) dz = \int_0^{\infty} V'(f(k)z) f'(k) z \psi(z) dz.$$

□

Now the proof of the Ramsey–Euler equation can be completed.

Proof of Proposition 3.2. Evidently $\sigma(y)$ solves

$$u'(y-k) - \varrho \frac{d}{dk} \int V[f(k)z]\psi(z)dz = 0.$$

The result now follows from Lemma 5.4, given that $V'(y) = u'(y - \sigma(y))$ Lebesgue almost everywhere. □

5.3. **Dynamics.** In the following discussion let an optimal policy σ be given. We simplify notation by defining the map S by $S(y) := f(\sigma(y))$. The most important properties of S are that—when σ is optimal—S is monotone nondecreasing and $S(y) = 0 \implies y = 0$ (Lemma 3.2 and Proposition 3.1, Part 1).

Let $D := \{g \in L_1(\mathbb{R}_+) : g \geq 0, \int g = 1\}$ be the set of density functions on $\mathbb{R}_+$. In general, D will be given the relative topology from the L_1 norm topology. In the sequel our notation does not distinguish between a distribution $\varphi \in \mathscr{P}$ and its density function in D. For example, if $\varphi \in \mathscr{P}$, the statement $\varphi \in D$ means that φ is absolutely continuous with respect to

[14]In particular, h' is bounded on compact sets, because $h'(x) = V'(e^x)e^x$, and $V'(y) = u'(y - \sigma(y))$ when it exists (i.e., when the function V' is not set to zero). The latter is bounded on compact sets by Lemma 5.3. Also, V is absolutely continuous because countinuous functions of bounded variation (provided by monotonicity here) fail to be absolutely continuous only if they have infinite derivative on an uncountable set (Saks, 1937, p. 128). This is impossible by Proposition 3.1, Part 4.

Lebesgue measure and can be represented by a density, which is also denoted φ.

Since S is zero only at zero, we can define the so-called Markov operator $P\colon L_1(\mathbb{R}_+) \ni g \mapsto Pg \in L_1(\mathbb{R}_+)$ by

$$(Pg)(y') = \int k(y, y')g(y)dy, \tag{17}$$

where

$$k(y, y') := \psi\left(\frac{y'}{S(y)}\right)\frac{1}{S(y)}. \tag{18}$$

The importance of the Markov operator is that for our model it generates by iteration the sequence of marginal densities (φ_t) for the Markov chain (y_t), $y_{t+1} = f(\sigma(y_t))\varepsilon_t$.

In the following, let P^t mean t compositions of P with itself. Also note that P maps D into itself, as is easily shown by Fubini's theorem.

Lemma 5.5. *If φ_0 is any initial condition, then $\varphi_1 \in D$ and $\varphi_t = P^{t-1}\varphi_1$ for all $t \geq 2$. Also, if $\varphi_0 \in D$, then $\varphi_t = P^t\varphi_0$ for all $t \geq 1$.*

Proof. Since by assumption $\varphi_0(\{0\}) = 0$, it is easy to see from (4) that $\varphi_1 \in D$. Now if $\varphi_t \in D$, then using (4), (17), (18) and the change of variable $y' = S(y)z$ gives $\varphi_{t+1} = P\varphi_t$, which completes the proof of the first statement. That the second statement is true follows from the same kind of argument[15] □

Corollary 5.1. *If an equilibrium φ^* exists then it is in D.*

Proof. Take $\varphi_0 = \varphi^*$ and apply the lemma. □

The next lemma is just translating the definitions of stability and sweeping given above—which have been formulated to fit in with the stochastic growth literature—to the language of Markov operators, where standard results are available.

[15] For more details on Markov operators see for example the monograph of Lasota and Mackey (1994). For a previous application in economics see Stachurski (2002).

Lemma 5.6. *Let σ be a fixed optimal policy, and let P be the corresponding Markov operator.*

1. *The economy is globally stable in the sense of Definition 4.1 if and only if there is a unique $\varphi^* \in D$ with $P\varphi^* = \varphi^*$ and $P^t\varphi \to \varphi^*$ in L_1 as $t \to \infty$ for every $\varphi \in D$.*
2. *The economy is globally collapsing to the origin in the sense of Definition 4.2 if and only if $\int_a^\infty P^t\varphi(y)dy \to 0$ for every $\varphi \in D$ and every $a > 0$.*

Proof. (Part 1, $\Rightarrow$) If $\varphi \in D$ then by hypothesis the trajectory (φ_t) starting at $\varphi_0 = \varphi$ converges to an equilibrium $\varphi^* \in \mathscr{P}$, and since $\varphi_t = P^t\varphi$ we have $P^t\varphi \to \varphi^*$. As D is complete we must have $\varphi^* \in D$, and in fact $P\varphi^* = \varphi^*$ by L_1-continuity of P (Lasota and Mackey 1994, Prop 3.1.1). If P has another fixed point in D, then it is easy to check that this fixed point satisfies (5), which contradicts uniqueness of equilibrium.

(Part 1, $\Leftarrow$) If P has a fixed point in $\varphi^* \in D$, then φ^* satisfies (5) as above, and hence is an equilibrium for the economy in $\mathscr{P}$. If φ^{**} is another equilibrium in $\mathscr{P}$, then $\varphi^{**} \in D$ by Corollary 5.1, and hence $P\varphi^{**} = \varphi^{**}$ as is easily verified from (5), contradicting uniqueness. If $\varphi_0 \in \mathscr{P}$ is any initial condition, then since $\varphi_1 \in D$ and $\varphi_t = P^{t-1}\varphi_1 \to \varphi^*$, we have convergence to the equilibrium from every initial condition.

The proof of Part 2 is a similar definition chasing exercise. □

Proof of Proposition 4.1. Let P be the Markov operator corresponding to σ, and let k be as in (18). Consider the following two conditions:

(i) $P\varphi > 0$ a.e., $\forall\, \varphi \in D$.

(ii) $\forall \hat{y} > 0$, $\exists \varepsilon > 0$ and $\eta \geq 0$ with $\int \eta(x)dx > 0$ and

$$k(y, y') \geq \eta(y')\mathbf{1}_{(\hat{y}-\varepsilon, \hat{y}+\varepsilon)}(y), \quad \forall\, y, y'.$$

By Rudnicki (1995, Theorem 2 and Corollary 3), (i) and (ii) imply the the Foguel Alternative; in particular that either P has a unique fixed point $\varphi^* \in D$ and $P^t\varphi \to \varphi^*$ in L_1 for all $\varphi \in D$, or alternatively P is sweeping with respect to the compact sets, so that $\lim_{t\to\infty} \int_a^b P^t\varphi(y)dy = 0$ for any

$\varphi \in D$ and any $0 < a < b < \infty$. In the light of Lemma 5.6, then, to prove Proposition 4.1 it is sufficient to check (i), (ii) and, in addition,

$$\lim_{b \to \infty} \limsup_{t \to \infty} \int_b^\infty P^t \varphi(y) dy = 0, \quad \forall \varphi \in D, \tag{19}$$

where (19) demonstrates that sweeping occurs not just with respect to any interval $[a, b]$, $a > 0$, but in fact to any interval $[a, \infty)$.

Condition (i) is immediate from the assumption that ψ is everywhere positive, in light of (17) and (18). Regarding condition (ii), pick any $\hat{y} > 0$ and any ε such that $\hat{y} - \varepsilon > 0$. Also let $0 < \gamma_0 < \gamma_1 < \infty$. Define

$$\delta_0 := \frac{\gamma_0}{S(\hat{y} + \varepsilon)}, \quad \delta_1 := \frac{\gamma_1}{S(\hat{y} - \varepsilon)}.$$

Note that $\inf_{z \in [\delta_0, \delta_1]} \psi(z) > 0$ by (S3) and strict positivity. Set

$$r := \frac{\inf_{z \in [\delta_0, \delta_1]} \psi(z)}{S(\hat{y} + \varepsilon)}, \quad \eta := r \mathbf{1}_{[\gamma_0, \gamma_1]}.$$

Then η has the required properties.

Regarding (19), from (F2) there exists a $\alpha \in (0, 1)$ and $m < \infty$ be such that $S(y) \leq \alpha y + m$ for all $y \in \mathbb{R}_+$. Then

$$y_{t+1} \leq (\alpha y_t + m)\varepsilon_t. \tag{20}$$

Since y_t and ε_t are independent and $\mathbb{E}^{\mathbf{P}} \varepsilon = 1$ we have

$$\mathbb{E}^{\mathbf{P}} y_{t+1} \leq \alpha \mathbb{E}^{\mathbf{P}} y_t + m. \tag{21}$$

Using an induction argument gives

$$\mathbb{E}^{\mathbf{P}} y_t \leq \alpha^t \mathbb{E}^{\mathbf{P}} y_0 + (1 + \alpha + \cdots + \alpha^{t-1}) m \leq \alpha^t \mathbb{E}^{\mathbf{P}} y_0 + \frac{m}{1 - \alpha}. \tag{22}$$

Suppose that $\mathbb{E}^{\mathbf{P}} y_0 < \infty$. Then from (22) it follows that

$$\limsup_{t \to \infty} \mathbb{E}^{\mathbf{P}} y_t \leq \frac{m}{1 - \alpha}. \tag{23}$$

By the Chebychev inequality, $\int_b^\infty P^t \varphi(y) dy \leq \mathbb{E}^{\mathbf{P}} y_t b^{-1}$. From (23) it then follows that (19) holds for all φ with $\mathbb{E}^{\mathbf{P}} y_0 := \int y \varphi(y)\, dy < \infty$. This set (all densities with finite first moments) is norm-dense in D, and P is an L_1 contraction. Together, these facts imply that condition (19) in fact holds for every $\varphi \in D$ (Lasota and Mackey 1994, p. 126). □

Proof of Lemma 4.1. By Corollary 3.1, any pair of optimal policies is equal almost everywhere. Inspection of (18) and (17) indicates that they will have identical Markov operators. Part 1 now follows from Lemma 5.6. The proof of Part 2 is similar. □

Proof of Lemma 4.2. Regarding Part 1, let σ_0 (resp. σ_1) be an optimal policy for E_0 (resp. E_1), let P_0 and P_1 be the corresponding Markov operators—defined by (17) and (18)—and let $(y_t^0)_{t \geq 0}$ and $(y_t^1)_{t \geq 0}$ be the respective income processes. By Lemmas 4.1 and 5.6 it is sufficient to show that for any $\varphi \in D$ and any $a > 0$ we have

$$\lim_{t \to \infty} \int_a^\infty P_0^t \varphi(y) dy = 0. \tag{24}$$

From Lemma 3.3 we have $\sigma_1 \geq \sigma_0$ pointwise on $\mathbb{R}_+$, so it is clear (by induction) that

$$y_t^1 \geq y_t^0 \text{ pointwise on } \Omega \text{ for any } t.$$

$$\therefore \quad \{y_t^0 \geq a\} \subset \{y_t^1 \geq a\}.$$

$$\therefore \quad \int_a^\infty P_0^t \varphi(y) dy = \mathbf{P}\{y_t^0 \geq a\} \leq \mathbf{P}\{y_t^1 \geq a\} = \int_a^\infty P_1^t \varphi(y) dy.$$

By Lemma 5.6 and the hypothesis, the right hand side converges to zero as $t \to \infty$, which proves (24). □

Proof of Proposition 4.3. For this proof we set $x_t := \ln y_t$, and define $\eta := \ln \varepsilon - \alpha$ and $T \colon \mathbb{R} \ni x \to \ln f(\sigma(e^x)) + \alpha$, so that $x_{t+1} = T(x_t) + \eta_t$, where $\mathbb{E}^{\mathbf{P}} \eta_t = 0$.

(Part 1) By the condition, $\limsup_{x \to -\infty} (T(x) - x) < 0$, implying the existence of an $m \in \mathbb{R}$ and $a > 0$ such that $T(x) \leq x - 2a$, for all $x \leq m$.

$$\therefore \quad x_{t+1} \leq x_t + \eta_t - 2a, \quad \forall x_t \leq m.$$

Let $\hat{x}_t := x_t - m$ and $\hat{\eta}_t := \eta_t - a$. Then

$$\hat{x}_{t+1} \leq \hat{x}_t + \hat{\eta}_t - a, \quad \forall \hat{x}_t \leq 0. \tag{25}$$

Define $\Omega_0 := \{\omega \in \Omega \,:\, \sup_{T \geq 0} \sum_{t=0}^T \hat{\eta}_t(\omega) \leq 0\}$. Since $\mathbb{E}^{\mathbf{P}} \hat{\eta}_t = -a < 0$, it follows that $\mathbf{P}(\Omega_0) > 0$ (Borovkov, 1999—see the discussion of factorization identities). From (25) we have

$$\hat{x}_t \leq \hat{x}_0 + \hat{\eta}_0 + \cdots + \hat{\eta}_{t-1} - ta \;\text{ for } \omega \in \Omega_0,$$

so if $\mathbf{P}\{\hat{x}_0 \leq 0\} = 1$, then $\mathbf{P}\{x_t \leq -at\} \geq \mathbf{P}(\Omega_0) > 0$ for all t. Since $\{\hat{x}_t \leq -at\} = \{y_t \leq e^{m-at}\}$, we have shown the existence of an initial condition y_0 ($\mathbf{P}\{\hat{x}_0 \leq 0\} = 1$ if y_0 is chosen s.t. $\mathbf{P}\{y_0 \leq e^m\} = 1$) with the property

$$\liminf_{t\to\infty} \mathbf{P}\{y_t \leq c\} = \liminf_{t\to\infty} \varphi_t([0,c]) \geq \mathbf{P}(\Omega_0) > 0.$$

But then φ_t cannot converge in L_1 to any $\varphi^* \in D$. (Elements of D are the only candidates for equilibria by Corollary 5.1. If $\varphi_t \to \varphi^* \in D$ then $\varphi_t([0,c]) \to \varphi^*([0,c])$, so choosing $c > 0$ such that $\varphi^*([0,c]) < \mathbf{P}(\Omega_0)$ leads to a contradiction.) Therefore the economy is not globally stable, and it follows from Proposition 4.1 that it must be collapsing to the origin.

(Part 2) By the condition, $\liminf_{x\to-\infty}(T(x) - x) > 0$, so there is an $m \in \mathbb{R}$ and $a > 0$ such that $T(x) \geq x + a$ whenever $x \leq m$. Let $\hat{x} := x - m$ and $\hat{\eta} := \eta + a$. Then $\hat{x}_{t+1} \geq \hat{x}_t + \hat{\eta}_t$ whenever $\hat{x}_t \leq 0$. Also, since T is nondecreasing, $\hat{x} \geq 0$ implies $T(x) \geq m + a$. Therefore $\hat{x}_t \geq 0 \implies \hat{x}_{t+1} \geq \hat{\eta}_t$.

$$\therefore \quad \hat{x}_{t+1} \geq \hat{x}_t^- + \hat{\eta}_t \geq (\hat{x}_t^- + \hat{\eta}_t)^-, \tag{26}$$

where we have introduced the notation $x^- = \min(0,x)$, and also $x^+ = \max(0,x)$.

Assume to the contrary that the economy is not globally stable, in which case it must be sweeping from the sets $[c,\infty)$, $c > 0$, so that for each $c \in \mathbb{R}$ we have

$$\lim_{t\to\infty} \mathbf{P}\{\hat{x}_t \leq c\} = 1. \tag{27}$$

Let us introduce now the process (z_t) defined by $z_0 := \hat{x}_0^-$, $z_{t+1} := (z_t + \hat{\eta}_t)^-$. By (26) we have $z_t \leq \hat{x}_t$ for all t. Since $\hat{\eta}_0$ is integrable, there is an $L \in \mathbb{R}$ such that $\mathbb{E}^{\mathbf{P}}(\hat{\eta}_0 - L)^+ < a/3$. Let y_0 be chosen so that $\hat{x}_0$ is also integrable. Then $\mathbb{E}^{\mathbf{P}}|z_0| < \infty$, and in fact $\mathbb{E}^{\mathbf{P}}|z_t| < \infty$ for all t. From (27) and $z_t \leq \hat{x}_t$ we have

$$\lim_{t\to\infty} \mathbf{P}\{z_t \leq -L\} = 1.$$

Choose t_0 so that $\mathbf{P}\{z_t > -L\} < a/(3L)$ when $t \geq t_0$. Since $z_t \leq 0$, then, $t \geq t_0$ implies $\mathbb{E}^{\mathbf{P}}(z_t + L)^+ < a/3$. Therefore,

$$\begin{aligned}\mathbb{E}^{\mathbf{P}} z_{t+1} = \mathbb{E}^{\mathbf{P}}(z_t + \hat{\eta}_t)^- &= \mathbb{E}^{\mathbf{P}}(z_t + \hat{\eta}_t) + \mathbb{E}^{\mathbf{P}}(z_t + \hat{\eta}_t)^+ \\ &\geq \mathbb{E}^{\mathbf{P}} z_t - \mathbb{E}^{\mathbf{P}}\hat{\eta}_t - \mathbb{E}^{\mathbf{P}}(z_t + L)^+ - \mathbb{E}^{\mathbf{P}}(\hat{\eta}_t - L)^+ \\ &> \mathbb{E}^{\mathbf{P}} z_t + \frac{a}{3},\end{aligned}$$

which contradicts $z_t \leq 0$ for all t. □

Proof of Proposition 4.4. By the Portmanteau Theorem (Shiryaev, 1996, Theorem III.1.1), $\varphi_n^* \to \delta_0$ weakly if and only if

$$\liminf_{n\to\infty} \varphi_n^*(G) \geq \delta_0(G) \quad \text{for every open set } G \subset \mathbb{R}_+.$$

Here by "open" we refer of course to the relative topology on $\mathbb{R}_+$. Evidently the above condition is equivalent to $\lim_n \varphi_n^*(G) = 1$ for all open G containing 0, which in turn is equivalent to

$$\lim_{n\to\infty} \varphi_n^*([a,\infty)) = 0, \quad \forall a > 0.$$

Take (σ_n) to be any sequence of optimal policies corresponding to $\varrho_n \to 0$. Let (y_t^n) be the Markov chain generated by σ_n and fixed initial distribution $y_0 \sim \varphi_0$ (i.e., $y_{t+1}^n = f(\sigma_n(y_t^n))\varepsilon_t$). Here $y_0 = y_0^n$ is chosen so that $\mathbb{E}^{\mathbf{P}} y_0 < \infty$.

Consider the probability that y_t^n exceeds a. For each real R we have

$$\begin{aligned}\text{(28)} \quad \mathbf{P}\{y_t^n \geq a\} = \mathbf{P}(\{y_t^n \geq a\} \cap \{y_{t-1}^n \leq R\}) \\ + \mathbf{P}(\{y_t^n \geq a\} \cap \{y_{t-1}^n > R\}).\end{aligned}$$

Consider the second term. We claim that

$$\text{(29)} \qquad \forall r > 0,\ \exists R \in \mathbb{R} \text{ s.t. } \sup_{n\in\mathbb{N}} \sup_{t\geq 0} \mathbf{P}\{y_t^n > R\} < r.$$

To see this, fix $r > 0$, and pick any $n \in \mathbb{N}$. Define a sequence (ξ_t) of random variables on $(\Omega, \mathscr{F}, \mathbf{P})$ by $\xi_0 = y_0$, $\xi_{t+1} = (\alpha\xi_t + \beta)\varepsilon_t$, where $y \mapsto \alpha y + \beta$ is an affine function dominating f on $\mathbb{R}_+$ and satisfying $\alpha < 1$ (see the comment after Assumption 2.2). From the definition of y_t^n, the fact that $\sigma_n(y) \leq y$ and $f(y) \leq \alpha y + \beta$, it is clear that $y_t^n \leq \xi_t$ pointwise on Ω for all t, and hence

$$\forall R \in \mathbb{R}, \quad \{y_t^n > R\} \subset \{\xi_t > R\}.$$

$$\therefore \quad \mathbf{P}\{y_t^n > R\} \leq \mathbf{P}\{\xi_t > R\}, \quad \forall t \geq 0. \tag{30}$$

Since ξ_t and ε_t are independent, $\mathbb{E}^{\mathbf{P}}\xi_{t+1} = \alpha\mathbb{E}^{\mathbf{P}}\xi_t + \beta$. It follows that

$$\mathbb{E}^{\mathbf{P}}\xi_t \leq \alpha^t \mathbb{E}^{\mathbf{P}}\xi_0 + \frac{\beta}{1-\alpha} \leq \mathbb{E}^{\mathbf{P}}\xi_0 + \frac{\beta}{1-\alpha}$$

for all t. Since $\mathbb{E}^{\mathbf{P}}\xi_0 = \mathbb{E}^{\mathbf{P}}y_0 < \infty$ we see that $\mathbb{E}^{\mathbf{P}}\xi_t \leq C$ for all t, where C is a finite constant. By the Chebychev inequality, then,

$$\mathbf{P}\{\xi_t > R\} \leq \frac{\mathbb{E}^{\mathbf{P}}\xi_t}{R} \leq \frac{C}{R}, \quad \forall t \geq 0. \tag{31}$$

Combining (30) and (31) gives $\mathbf{P}\{y_t^n > R\} < C/R$ for all t and n. Since R is arbitrary the claim (29) is established.

Our objective was to bound the second term in (28). So fix $r > 0$. By (29) we can choose R so large that

$$\mathbf{P}\{y_t^n \geq a\} = \mathbf{P}(\{y_t^n \geq a\} \cap \{y_{t-1}^n \leq R\}) + \frac{r}{2} \tag{32}$$

for all t and all n. It remains to bound the first term. Let $(\varphi_t^n) \subset \mathscr{P}$ be the sequence of marginal distributions associated with (y_t^n). From the well-known expression for the finite dimensional distribution of Markov chains on measurable rectangles (e.g., Shiryaev, 1996, Theorem II.9.2) we have

$$\mathbf{P}(\{y_t^n \geq a\} \cap \{y_{t-1}^n \leq R\}) = \int_0^R \int_a^\infty \psi\left(\frac{y'}{f(\sigma_n(y))}\right) \frac{1}{f(\sigma_n(y))} dy' \varphi_{t-1}(y) dy.$$

A change of variable gives

$$\int_a^\infty \psi\left(\frac{y'}{f(\sigma_n(y))}\right) \frac{1}{f(\sigma_n(y))} dy' = \psi([a/f(\sigma_n(y)), \infty)),$$

where, as always, we are using ψ to denote both the density and the measure $\varphi(dz) = \psi(z)dz$. From the proof of Lemma 3.4, we know that σ_n is dominated by an increasing function b_n which converges pointwise to zero. Therefore $f \circ \sigma_n$ is dominated by $f \circ b_n$, again an increasing function, which must by continuity of f converge pointwise and hence uniformly to zero on $[0, R]$. Combining this with the fact that $a > 0$ and ψ is a finite measure, there is an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$\psi([a/f(\sigma_n(y)), \infty)) < \frac{r}{2}, \quad \forall y \in [0, R].$$

But then

$$\mathbf{P}(\{y_t^n \geq a\} \cap \{y_{t-1}^n \leq R\}) \leq \int_0^R \frac{r}{2}\varphi_{t-1}(y)dy \leq \frac{r}{2}.$$

Using this inequality together with (28) and (32), we conclude that for all $r > 0$ there is an $N \in \mathbb{N}$ such that $n \geq N$ and $t \geq 0$ implies $\mathbf{P}\{y_t^n \geq a\} = \varphi_t^n([a, \infty)) < r$. Since $\varphi_t^n \to \varphi_n^*$ in L_1 as $t \to \infty$ and hence weakly in the sense of the topology induced on L_1 by L_∞, it follows that $\varphi_t^n([a, \infty)) \to \varphi_n^*([a, \infty))$ in $\mathbb{R}$ as $t \to \infty$, so that $\varphi_n^*([a, \infty)) \leq r$ is also true. That is, $\lim_{n\to\infty} \varphi_n^*([a, \infty)) = 0$, as was to be proved. □

Appendix A

First we need the following lemma regarding continuity of translations in L_1, which is well-known.

Lemma A.1. *Let g be in $L_1(\mathbb{R})$. If $\tau(t) := \|g(x - t) - g(x)\|$, then τ is bounded on $\mathbb{R}$, and $\tau(t) \to 0$ as $t \to 0$.*

Now define the real number $\mu'(r)$ to be $-\int h(x+r)g'(x)\,dx$, which is clearly finite. By the Fundamental Theorem of Calculus,

$$\begin{aligned}\mu(r + t) - \mu(r) - \mu'(r)t &= \int h(x + r)(g(x - t) - g(x) + g'(x)t)\,dx \\ &= -t \int h(x + r) \int_0^1 (g'(x - ut) - g'(x))\,du\,dx.\end{aligned}$$

Taking absolute values, using (ii) and Fubini's theorem,

$$\left|\frac{\mu(r + t) - \mu(r)}{t} - \mu'(r)\right| \leq M \int_0^1 \int |g'(x - ut) - g'(x)|\,dx\,du \tag{33}$$

for some M. By Lemma A.1, $\int |g'(x - ut) - g'(x)|\,dx$ is uniformly bounded in u and converges to zero as $t \to 0$ for each $u \in [0, 1]$. By Lebesgue's Dominated Convergence Theorem the term on the right hand side of (33) then goes to zero and

$$\mu'(r) = -\int h(x + r)g'(x)\,dx$$

as claimed.

Regarding continuity of the derivative, we have

$$|\mu'(r+t) - \mu'(r)| \leq \int h(x)|g'(x-r-t) - g'(x-r)|\, dx$$
$$\leq M \int |g'(x-t) - g'(x)|\, dx.$$

Continuity now follows from Lemma A.1.

Next we argue that under (iii)–(v),

$$\mu'(r) = \int h'(x+r) g(x)\, dx \tag{34}$$

is also valid. To begin, define $\mu'_h(r)$ to be the right hand side of (34). This number exists in $\mathbb{R}$, because

$$h'(x+r) = \liminf_{t \downarrow 0} \frac{h(x+r+t) - h(x+r)}{t}$$

almost everywhere by either (iii) or (iv), and hence

$$\mu'_h(r) = \int \liminf_{t \downarrow 0} \frac{h(x+r+t) - h(x+r)}{t} g(x) dx$$
$$\leq \liminf_{t \downarrow 0} \int \frac{h(x+r+t) - h(x+r)}{t} g(x) dx = \mu'(r).$$

Here the inequality follows from the assumption that h is increasing, which gives nonnegativity of the difference quotient, and Fatou's Lemma.

By (iv) the Fundamental Theorem of Calculus applies to h, and

$$\mu(r+t) - \mu(r) - \mu'_h(r)t = \int (h(x+t) - h(x) - h'(x)t) g(x-r)\, dx$$
$$= t \int \int_0^1 (h'(x+ut) - h'(x)) g(x-r)\, dx\, du.$$

Some simple manipulation gives

$$\mu'_h(r) = \mu'(r) - \lim_{t \to 0} \int \int_0^1 (h'(x+ut) - h'(x)) g(x-r)\, dx\, du.$$

Thus it is sufficient to now show that

$$\lim_{t \to 0} \int_0^1 \int |h'(x+ut) - h'(x)| g(x-r)\, dx\, du = 0.$$

The inner integral is bounded independent of u, because it is less than

$$\int h'(x+ut) g(x-r)\, dx + \int h'(x) g(x-r)\, dx \leq \mu'(r+ut) + \mu'(r),$$

which is bounded for $u \in [0,1]$ by continuity of μ'. Thus by Lebesgue's Dominated Convergence Theorem we need only prove that

$$\lim_{t\to 0} \int |h'(x+ut) - h'(x)| g(x-r)\, dx = 0.$$

Adding and subtracting appropriately, this integral is seen to be less than

$$\int |h'(x+ut)g(x-r+ut) - h'(x)g(x-r)| dx + \int |h'(x+ut)g(x-r) - h'(x+ut)g(x-r+ut)|\, dx. \tag{35}$$

Consider the first integral in the sum. By Lemma A.1, we can choose a $\delta_0 > 0$ such that $|t| \le \delta_0$ implies

$$\int |h'(x+ut)g(x-r+ut) - h'(x)g(x-r)| dx < \frac{\varepsilon}{3}.$$

The second integral in the sum can be written as

$$\int_{|x|\le R} |h'(x+ut)g(x-r) - h'(x+ut)g(x-r+ut)|\, dx + \int_{|x|\ge R} |h'(x+ut)g(x-r) - h'(x+ut)g(x-r+ut)|\, dx.$$

By the usual property of L_1 functions, we can choose R such that the integral over $|x| \ge R$ is less than $\varepsilon/3$ for all t with $|t| \le \delta_0$.

To summarize the results so far, we have $|t| \le \delta_0$ implies

$$\int |h'(x+ut) - h'(x)| g(x-r)\, dx < \frac{2\varepsilon}{3} + \int_{|x|\le R} |h'(x+ut)g(x-r) - h'(x+ut)g(x-r+ut)|\, dx.$$

Finally, since h' is bounded on compact sets,

$$h'(x+ut) \le M, \quad \forall x, t \text{ with } |x| \le R,\ |t| \le \delta_0.$$

Therefore $|t| \le \delta_0$ implies

$$\int |h'(x+ut) - h'(x)| g(x-r)\, dx < \frac{2\varepsilon}{3} + M \int |g(x-r) - g(x-r+ut)|\, dx.$$

By Lemma A.1 there is a $\delta_1 > 0$ such that

$$M \int |g(x-r) - g(x-r+ut)|\, dx < \frac{\varepsilon}{3}$$

whenever $|t| < \delta_1$. Now setting $\delta := \delta_0 \wedge \delta_1$ gives

$$|t| \leq \delta \implies \int |h'(x + ut) - h'(x)| g(x - r)\, dx < \varepsilon$$

as required.

References

[1] Amir, R. (1997): A New Look at Optimal Growth Under Uncertainty, *Journal of Economic Dynamics and Control,* 22, 67–86.

[2] Amir, R., L.J. Mirman and W.R. Perkins (1991): One-Sector Nonclassical Optimal Growth: Optimality Conditions and Comparative Dynamics, *International Economic Review,* 32 (3), 625–644.

[3] Askri, K. and C. Le Van (1998): Differentiability of the Value Function of Nonclassical Optimal Growth Models, *Journal of Optimization Theory and Applications,* 97 (3), 591–604.

[4] Blume, L., D. Easley and M. O'Hara (1982): Characterization of Optimal Plans for Stochastic Dynamic Economies, *Journal of Economic Theory,* 28, 221–234.

[5] Borovkov, A.A. (1999): *Probability Theory*, Taylor and Francis.

[6] Brock, W. A. and L. Mirman (1972): Optimal Economic Growth and Uncertainty: The Discounted Case, *Journal of Economic Theory* 4, 479–513.

[7] Danthine, J-P. and J. B. Donaldson (1981): Stochastic Properties of Fast vs. Slow Growth Economies, *Econometrica,* 49 (4), 1007–1033.

[8] Dechert, W. D. and K. Nishimura (1983): A Complete Characterization of Optimal Growth Paths in an Aggregated Model with Non-Concave Production Function, *Journal of Economic Theory,* 31, 332–354.

[9] Dudley, R. M. (2002): *Real Analysis and Probability*, Cambridge Studies in Advanced Mathematics No. 74, Cambridge.

[10] Foguel, S. R. (1969): *The Ergodic Theory of Markov Processes,* Van Nostrand Reinhold, New York.

[11] Futia, C. A. (1982): Invariant Distributions and the Limiting Behavior of Markovian Economic Models, *Econometrica,* 50, 377–408.

[12] Hopenhayn, H. A. and E. C. Prescott (1992): Stochastic Monotonicity and Stationary Distributions for Dynamic Economies, *Econometrica,* 60, 1387–1406.

[13] Joshi, S. (1997): Turpike Theorems in Nonconvex Nonstationary Environments, *International Economic Review,* 38 (1), 225–248.

[14] Kamihigashi, T. (2003): Almost Sure Convergence to Zero in Stochastic Growth Models, mimeo, RIEB, Kobe University.

[15] Kamihigashi, T. and S. Roy (2003): A Non-smooth, Non-Convex Model of Economic Growth," manuscript.

[16] Kydland, F. and E. C. Prescott (1982): Time to Build and Aggregate Fluctuations, *Econometrica,* 50, 1345–1370.

[17] Lasota, A. and M. C. Mackey, 1994, *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics,* 2nd edition, Springer-Verlag, New York.

[18] Long, J. B. and C. I. Plosser (1983): Real Business Cycles, *Journal of Political Economy,* 91, 39–69.

[19] Lucas, R. E. (1986): Adaptive Behavior and Economic Theory, *The Journal of Business,* Vol. 59, No. 4, 385–399.

[20] Majumdar, M. and T. Mitra (1982): Intertemporal Allocation with Nonconvex Technology, *Journal of Economic Theory,* 27, 101–136.

[21] Majumdar, M., T. Mitra and Y. Nyarko (1989): Dynamic Optimization Under Uncertainty: Non-convex feasible set" in *Joan Robinson and Modern Economic Theory,* (Ed.) G. R. Feiwel, MacMillan Press, New York.

[22] Mitra, T. and S. Roy (2003): Optimal Exploitation of Resources under Uncertainty and the Extinction of Species, manuscript.

[23] Mirman, L. J. (1970): Two Essays on Uncertainty and Economics, Ph.D. Thesis, University of Rochester.

[24] Mirman, L. J., O. F. Morand and K. Reffett (2003): A Qualitative Approach to Markovian Equilibrium in Infinite Horizon Economies with Capital, manuscript.

[25] Mirman, L. J. and I. Zilcha (1975): On Optimal Growth under Uncertainty, *J. Econ. Theory* **11** (1975), 329–339.

[26] Prescott, E. C. (2002): Non-Convexities in Quantitative General Equilibrium Studies of Business Cycles, Staff Report No. 312, Federal Reserve Bank of Minneapolis.

[27] Rudnicki, R. (1995): On Asymptotic Stability and Sweeping for Markov Operators, *Bulletin of the Polish Academy of Science: Mathematics,* 43, 245–262.

[28] Saks, S. (1937): *Theory of the Integral*, Monografie Matematyezne, Warsaw.

[29] Schenk-Hoppé, K. R. (2002): Poverty Traps and Business Cycles in a Stochastic Overlapping Generations Economy with S-shaped Law of Motion. Discussion Paper No. 02-13, Institute of Economics, University of Copenhagen.

[30] Shiryaev, A. N. (1996) *Probability* Springer-Verlag, New York.

[31] Skiba, A. K. (1978): Optimal Growth with a Convex-Concave Production Function, *Econometrica,* 46, 527–539.

[32] Stachurski, J. (2002): Stochastic Optimal Growth with Unbounded Shock, *Journal of Economic Theory,* 106, 40–65.

[33] Stokey, N. L., R. E. Lucas and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics,* Harvard University Press, Massachusetts.

Institute of Economic Research, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

E-mail address: nishimura@kier.kyoto-u.ac.jp

Institute of Mathematics, Polish Academy of Sciences and Institute of Mathematics, Silesian University, Bankowa 14, 40–007 Katowice, Poland

E-mail address: rudnicki@us.edu.pl

Center for Operations Research and Econometrics, Université catholique de Louvain, 34 Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium

E-mail address: stachurski@core.ucl.ac.be