Stochastic Optimal Policies when the Discount Rate Vanishes ⋆

Kazuo Nishimura

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

John Stachurski *

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

and

Department of Economics, The University of Melbourne, VIC 3010, Australia

---

Abstract

Dutta [1991. Journal of Economic Theory 55, 64–94] showed that long-run optimality of the limit of discounted optima when the discount rate vanishes is implied by a certain bound on the value function of the optimal program. We introduce a new method to verify this bound using coupling techniques.

Journal of Economic Literature Classifications C60, C61

Key words: Dynamic programming, Long-run optimality

⋆ The authors thank the anonymous referees for suggestions which led to important extensions of our initial results. The second author is grateful for financial support from Australian Research Council Grant DP0557625.

* Corresponding author.

Email addresses: `nishimura@kier.kyoto-u.ac.jp` (Kazuo Nishimura), `john@kier.kyoto-u.ac.jp` (John Stachurski).

Preprint submitted to Elsevier Science *25 April 2006*

1 Introduction

Discounted dynamic programming is a standard paradigm for analyzing economic outcomes when expectations are rational and information is perfect. (For dynamics in imperfect information economies see, for example, Chiarella and Szidarovzky (2001) and references.) An established theory exists, along with practical methods of numerical computation. However, optimal behavior when the future is not discounted has also been studied, perhaps most famously in the classic paper of Ramsey (1928). [1] Another well-known example is the no-discounting paper by Brock and Mirman (1973), albeit much less so than its famous discounting cousin (1972).

A number of no-discounting criteria exist for optimality. In the mathematical literature on stochastic dynamic programming, however, no-discounting research is now mainly focused on long-run average reward (AR) optimality, which maximizes the average of the undiscounted sequence of period rewards. [2] For example, AR-optimalization is routinely applied to on-line computer task scheduling and network routing. [3]

It is of great practical interest to identify relationships between discounted reward (DR) optimal policies and AR-optimal policies. For example, if π_ϱ is a DR-optimal policy for discount factor $\varrho \in (0, 1)$, and if π_ϱ converges to a

[1] According to Ramsey, "discount[ing] later enjoyments in comparison with earlier ones [is] ethically indefensible, and arises merely from the weakness of the imagination" (1928, p. 543).

[2] If $(r_t)_{t \geq 0}$ is a bounded sequence of rewards, then the average is usually defined to be $\liminf_{t \to \infty} (1/t) \sum_{s=0}^{t-1} r_s$. For a recent treatmentment of AR optimality see the excellent paper of Jaśkiewicz and Nowak (2006).

[3] In economic growth another popular criterion for optimality is "overtaking," which requires that expected period reward eventually dominates that of other policies. This is closely related to so-called turnpike theory. For a survey of the literature see McKenzie (1998).

limit π_1 when $\varrho \uparrow 1$, it seems likely that π_1 will be—at least in some sense—long-run optimal. One would like to know under what conditions, if any, π_1 is AR-optimal.

An important contribution to our understanding of the relationship between DR- and AR-optimality is the study of Dutta (1991). He showed that when the pointwise limit π_1 exists it is AR-optimal, provided that the optimal program satisfies a certain "value boundedness" condition, which is stated in terms of the value function. We introduce a new method for verifying this condition, based on coupling techniques.

Coupling involves making statements about two probability distributions P and P' by setting up random elements X and X' on a common probability space, where X (resp., X') has marginal distribution P (resp., P'). In our case the random elements are sequences generated by the same optimal program, but having different initial conditions. Their distributions are used to calculate value functions. [4]

Two applications are given. The first is for economies satisfying the "monotone mixing" conditions of Stokey, Lucas and Prescott (1989) and Hopenhayn and Prescott (1992). The second verifies the conjecture that π_1 defined above is AR-optimal for a relatively general neoclassical stochastic optimal growth model.

[4] It has been said that coupling proofs are like jokes: Detailed explanation ruins them. Unfortunately our paper is no exception. In applying coupling techniques, a number of our ideas draw on the study of ergodicity in Rosenthal (2003).

2 Formulation of the Problem

Let A and S be well-behaved topological spaces.[5] Let Γ be a continuous, nonempty, compact valued correspondence from S to A, representing feasible choices for each state $x \in S$, and let

$$K := \{(x, a) \in S \times A : a \in \Gamma(x)\}.$$

Let $r \colon K \to \mathbb{R}$ be a bounded reward function which is jointly measurable on K, with $a \mapsto r(x, a)$ continuous on $\Gamma(x)$ for each fixed $x \in S$.[6] Finally, let $(\xi_t)_{t=0}^{\infty}$ be an independent and identically distributed collection of random variables on probability space $(\Omega, \mathscr{F}, \mathbb{P})$, all taking values in $(Z, \mathscr{Z})$ and having distribution ν, and let $h \colon K \times Z \to S$ be a jointly measurable function, which updates the state according to $x' = h(x, a, \xi)$.[7] Suppose that $a \mapsto h(x, a, z)$ is continuous on $\Gamma(x)$ for each $x \in S$ and $z \in Z$.

Define Π to be the set of all feasible policies, which are measurable functions $\pi \colon S \to A$ satisfying $\pi(x) \in \Gamma(x)$ for all $x \in S$. Each $\pi \in \Pi$ determines a Markov process $(x_t)_{t=0}^{\infty}$ for the state via

$$x_{t+1} = h(x_t, \pi(x_t), \xi_t), \quad x_0 \text{ given.} \tag{1}$$

For each $\varrho \in (0, 1)$ and each $\pi \in \Pi$ let

$$v_{\varrho}(x_0, \pi) := \mathbb{E}\left[\sum_{t=0}^{\infty} \varrho^t r(x_t, \pi(x_t))\right],$$

[5] It is sufficient that they be separable and completely metrizable. All G_δ subsets of $\mathbb{R}^n$ have this property.

[6] In this paper, measurability in reference to functions on topological spaces always refers to Borel measurability.

[7] Note there is little loss of generality in assuming that $(\xi_t)_{t=0}^{\infty}$ is IID. If, for example, $(\xi_t)_{t=0}^{\infty}$ is first order Markov, one can always rewrite the transitions in the form $x' = h(x, a, \zeta)$ where $(\zeta_t)_{t=0}^{\infty}$ is IID, by a suitable transformation of the function h and the state space S.

where $(x_t)_{t=0}^{\infty}$ is given by (1). The optimal investment problem is then to solve

$$\max_{\pi \in \Pi} v_{\varrho}(x_0, \pi). \tag{2}$$

A policy is called DR-ϱ-optimal if it solves (2) for all $x_0 \in S$. It is well-known that under the current assumptions at least one DR-ϱ-optimal policy exists. The value function v_{ϱ} is defined at x_0 by $\sup_{\pi \in \Pi} v_{\varrho}(x_0, \pi)$.

The other optimality criterion we consider is AR-optimality. A policy is called AR-optimal if it solves

$$\max_{\pi \in \Pi} \liminf_{t \to \infty} \mathbb{E} \left[\frac{1}{t} \sum_{s=0}^{t-1} r(x_s, \pi(x_s)) \right], \tag{3}$$

where again π determines the process $(x_t)_{t=0}^{\infty}$ via (1).

One of the most useful conditions for linking DR- and AR-optimality is value boundedness:

Definition 2.1 *The dynamic programming problem* (Γ, r, h, ν) *is called* value bounded *if there exists an* $x' \in S$*, a function* $m_1 \colon S \to \mathbb{R}$ *and a constant* $m_2 < \infty$ *such that*

$$m_1(x) \leq v_{\varrho}(x) - v_{\varrho}(x') \leq m_2, \quad \forall x \in S, \ \varrho \in (0, 1).$$

For a standard class of optimal programs, Dutta (1991) showed that any pointwise limit of DR-ϱ optimal policies is AR-optimal whenever value boundedness holds. [8]

3 Results

We now develop an inequality which has obvious application in determining when economies are value bounded. The inequality is given in Theorem 3.1 below.

[8] See also the paper of Sennott (1986).

To begin, let $(\xi'_t)_{t=0}^\infty$ be another sequence of IID random variables on the initial probability space $(\Omega, \mathscr{F}, \mathbb{P})$, again taking values in $(Z, \mathscr{Z})$ and having distribution ν. Assume that $(\xi'_t)_{t=0}^\infty$ and the original sequence $(\xi_t)_{t=0}^\infty$ are independent. Also, for fixed $\varrho \in (0,1)$, let π_ϱ be any DR-ϱ-optimal policy, and let $(x_t)_{t=0}^\infty$ and $(x'_t)_{t=0}^\infty$ be two sequences satisfying $x_{t+1} = h(x_t, \pi_\varrho(x_t), \xi_t)$ and $x'_{t+1} = h(x'_t, \pi_\varrho(x'_t), \xi'_t)$, and starting from x_0 and x'_0 respectively. Define the random variable

$$\tau_\varrho(x_0, x'_0) \; := \inf\{t \geq 0 : v_\varrho(x_t) \leq v_\varrho(x'_t)\}, \tag{4}$$

with the usual convention that $\inf \varnothing = \infty$. Thus, τ_ϱ is the first time that $v_\varrho(x_t)$ falls below $v_\varrho(x'_t)$. The relevance of this "swapping time" follows from

Theorem 3.1 *Let $x_0, x'_0 \in S$, and let $\varrho \in (0,1)$. If $\bar{r} := \sup_{x,a} r(x,a)$, then*

$$v_\varrho(x_0) - v_\varrho(x'_0) \leq 2\bar{r}\, \mathbb{E}\, \tau_\varrho(x_0, x'_0).$$

Remark 3.1 *If $v_\varrho(x_0) \leq v_\varrho(x'_0)$ then $\tau_\varrho(x_0, x'_0) \equiv 0$ and the bound holds trivially. The more interesting case is where $v_\varrho(x_0) > v_\varrho(x'_0)$. In this case the intuition is as follows: The function v_ϱ ranks points in the state space according to their value. If an economy starting at the higher value state x_0 is expected to move quickly into an area of the state space with lower value than an economy which started at x'_0 (i.e., if $v_\varrho(x_t) \leq v_\varrho(x'_t)$ is expected to occur for small t), then the relative advantage of starting at the higher value state x_0 cannot be too large.*

Remark 3.2 *In general, the easiest way to prove that $\mathbb{E}\, \tau_\varrho(x_0, x'_0)$ is finite is to show that $\mathbb{P}\{\tau_\varrho(x_0, x'_0) > t\}$ goes to zero quickly with t, in which case the tail of the distribution is light and the mean is small.*

PROOF. [Proof of Theorem 3.1] Since ϱ is fixed in this proof we omit to use it as a subscript. To begin, note that

$$\begin{aligned} v(x_0) &= \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t r(x_t, \pi(x_t)) + \mathbb{E}\sum_{t=\tau}^{\infty} \varrho^t r(x_t, \pi(x_t)) \\ &= \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t r(x_t, \pi(x_t)) + \mathbb{E}\, \varrho^\tau v(x_\tau). \end{aligned}$$

The intuitively plausible second step is given a formal justification in the appendix. A similar argument for $v(x_0')$ gives

$$v(x_0') = \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t r(x_t', \pi(x_t')) + \mathbb{E}\, \varrho^\tau v(x_\tau').$$

By the definition of τ we have $\mathbb{E}\, \varrho^\tau v(x_\tau) \leq \mathbb{E}\, \varrho^\tau v(x_\tau')$, so subtracting one equality from the other gives

$$\begin{aligned} v(x_0) - v(x_0') &\leq \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t r(x_t, \pi(x_t)) - \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t r(x_t', \pi(x_t')) \\ &\leq \mathbb{E}\sum_{t=0}^{\tau-1} \varrho^t |r(x_t, \pi(x_t)) - r(x_t', \pi(x_t'))| \leq \mathbb{E}\sum_{t=0}^{\tau-1} 2\bar{r}. \end{aligned}$$

The last term is just $2\bar{r}\mathbb{E}\tau$, so the proof is done.

Evidently Theorem 3.1 has applications to the problem of value boundedness. In particular, the following corollary holds:

Corollary 3.1 *If there exists an* $m\colon S \to \mathbb{R}$ *s.t.* $\mathbb{E}\tau_\varrho(x, y) \leq m(y)$, $\forall x \in S$, $\forall \varrho \in (0, 1)$, *then the dynamic program defined by* (Γ, r, h, ν) *is valued bounded.*

PROOF. Fix $x' \in S$. By Theorem 3.1 and the hypothesis we have

$$v_\varrho(x) - v_\varrho(x') \leq 2\bar{r}\mathbb{E}\tau_\varrho(x, x') \leq m_2, \quad \forall\, x \in S,\ \varrho \in (0, 1),$$

where $m_2 := 2\bar{r}m(x')$. By the same argument we have

$$v_\varrho(x') - v_\varrho(x) \leq 2\bar{r}\mathbb{E}\tau_\varrho(x', x) \leq -m_1(x), \quad \forall\, x \in S,\ \varrho \in (0, 1),$$

where $m_1(x) := -2\bar{r}m(x)$.

$$\therefore \quad m_1(x) \leq v_\varrho(x) - v_\varrho(x') \leq m_2, \quad \forall\, x \in S,\ \varrho \in (0, 1).$$

Remark 3.3 *We make one further remark on the general theory. Clearly* $|v_\varrho(x) - v_\varrho(x')| \leq 2\bar{r}/(1-\varrho)$ *holds for any* x *and* x', *where as before* $\bar{r} := \sup r$. *Therefore when establishing value boundedness one can always restrict attention to* $\varrho \in [\hat{\varrho}, 1)$ *for some fixed* $\hat{\varrho} \in (0,1)$.

4 Application: Monotone Mixing

Our first application concerns monotone dynamic programs which satisfying a well-known monotone mixing condition. The mixing condition was popularized by Stokey, Lucas and Prescott (1989, Assumption 12.1) and Hopenhayn and Prescott (1992, Theorem 2), who used it to study ergodicity.

In addition to the assumptions in Section 2, we suppose that

Assumption 4.1 *The state space* S *is an order interval* $[a, b]$ *of* $\mathbb{R}^n$ *with its usual pointwise ordering. The map* $x \mapsto h(x, \pi_\varrho(x), z)$ *is monotone increasing on* S *for all* $z \in Z$ *and* $\varrho \in (0,1)$. *The map* $x \mapsto r(x, \pi_\varrho(x))$ *is also increasing for each* $\varrho \in (0,1)$.

Here for each $\varrho \in (0,1)$, π_ϱ is a corresponding optimal policy. Although Assumption 4.1 concerns optimal policies—which are not primitives of the model—conditions for monotonicity of optimal policies have been extensively investigated, so we do not persue the matter here. [9]

One can easily verify from Assumption 4.1 that v_ϱ is monotone increasing on S. We use this fact in the proofs without further comment.

The next assumption is the monotone mixing condition:

Assumption 4.2 *There exists an* $\varepsilon > 0$, *a* $c \in S$ *and an* $N \in \mathbb{N}$ *such that for all* $\varrho \in (0,1)$ *we have*

$$\mathbb{P}\{x_{t+N} \geq c \,|\, x_t = a\} \geq \varepsilon \text{ and } \mathbb{P}\{x_{t+N} \leq c \,|\, x_t = b\} \geq \varepsilon.$$

[9] See, for example, Hopenhayn and Prescott (1992), or Mirman, Morand and Reffett (2005) and references.

Combined with Assumption 4.1, Assumption 4.2 says that for any $x \in S$, both $\mathbb{P}\{x_{t+N} \geq c \,|\, x_t = x\}$ and $\mathbb{P}\{x_{t+N} \leq c \,|\, x_t = x\}$ exceed ε, and ε does not depend on ϱ.[10]

Remark 4.1 *In view of Remark 3.3, Assumption 4.2 need only hold for ϱ in a neighborhood of 1 .*

Theorem 4.1 *If Assumptions 4.1 and 4.2 both hold, then the dynamic program (Γ, r, h, ν) is value bounded.*

The intuition is straightforward. Pick initial conditions x and x'. By Corollary 3.1, it is sufficient to show that $\mathbb{E}\tau_\varrho(x, x')$ is bounded above by a finite constant which is independent of x and ϱ. By Assumption 4.1 the value function is increasing for all ϱ. Thus, if $x_t \leq x'_t$ then $\tau_\varrho(x, x') \leq t$. Every N steps, $(x_t)_{t=0}^\infty$ has an at least ε probability of entering $[a, c]$, as does $(x'_t)_{t=0}^\infty$ for the set $[c, b]$. By independence, both events occur simultaneously with probabilty ε^2. If they do, then $x_t \leq x'_t$. Thus, every N steps, there is an ε^2 chance that $v_\varrho(x_t) \leq v_\varrho(x'_t)$, which suggests the bound

$$\mathbb{P}\{\tau_\varrho(x, x') > kN\} \leq (1 - \varepsilon^2)^k, \quad \forall k \geq 0. \tag{5}$$

This rate of decrease is sufficient for $\mathbb{E}\tau_\varrho$ to be finite. In particular, since $\mathbb{P}\{\tau_\varrho(x, x') > t\}$ is decreasing in t we get

$$\begin{aligned}
\mathbb{E}\tau_\varrho = \sum_{t=1}^\infty t\mathbb{P}\{\tau_\varrho = t\} &\leq \sum_{t=1}^\infty t\mathbb{P}\{\tau_\varrho \geq t\} \\
&= \sum_{t=1}^\infty t\mathbb{P}\{\tau_\varrho > t - 1\} \\
&\leq N \sum_{k=1}^\infty kN\mathbb{P}\{\tau_\varrho > (k-1)N\},
\end{aligned}$$

which is dominated by $N^2 \sum_k k(1 - \varepsilon^2)^{(k-1)} < \infty$ as a result of (5).

[10] Conditional expectations of the form $\mathbb{P}\{x_{t+N} \in B \,|\, x_t = x\}$ are interpreted as the cannonical conditional expectation associated with $x \mapsto h(x, \pi_\varrho(x), z)$ and ν, and as such are uniquely defined. See the discussion of Markov kernels later in this section.

It remains to give a formal justification for (5). Some new notation will be helpful. Recall that a Markov kernel on topological space T is a map $\mathbf{N}\colon T \times \mathscr{B}(T) \to [0,1]$, where $\mathscr{B}(T)$ is the Borel sets on T, and $\mathbf{N}$ has the following properties: $x \mapsto \mathbf{N}(x, B)$ is Borel measurable for all $B \in \mathscr{B}(T)$, and $B \mapsto \mathbf{N}(x, B)$ is a Borel probability measure for all $x \in T$. One defines iterates of $\mathbf{N}$ by setting $\mathbf{N}^t(x, B) := \int \mathbf{N}(x, dy)\mathbf{N}^{t-1}(y, B)$. Intuitively, $\mathbf{N}^t(x, B)$ is the probability of the state moving from x now into set B t periods hence.

A sequence of T-valued random variables $(y_t)_{t=0}^{\infty}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ adapted to filtration $(\mathscr{F})_{t=0}^{\infty}$ is called a Markov process with Markov kernel $\mathbf{N}$ if $\mathbb{P}\{y_{t+1} \in B \,|\, \mathscr{F}_t\} = \mathbf{N}(y_t, B)$, for all $B \in \mathscr{B}(T)$ and all $t \geq 0$. In this case one can establish that $\mathbb{P}\{y_{t+j} \in B \,|\, \mathscr{F}_t\} = \mathbf{N}^j(y_t, B)$ for any $j \in \mathbb{N}$, and in fact the same is true if we replace t with a stopping time.

For both our processes $(x_t)_{t=0}^{\infty}$ and $(x'_t)_{t=0}^{\infty}$ the relevent kernels are given by $\mathbf{M}_\varrho(x, B) = \mathbb{P}\{h(x, \pi_\varrho(x), \xi_t) \in B\}$. The kernels are the same because ξ_t and ξ'_t both have identical distribution ν. The joint process $(x_t, x'_t)_{t=0}^{\infty}$ is Markov on $S \times S$, and in what follows $(\mathscr{F}_t)_{t=0}^{\infty}$ is always the natural filtration for this process. If $\hat{\mathbf{M}}_\varrho$ is the Markov kernel for $(x_t, x'_t)_{t=0}^{\infty}$, then independence of $(x_t)_{t=0}^{\infty}$ and $(x'_t)_{t=0}^{\infty}$ implies that

$$\hat{\mathbf{M}}_\varrho(x, x', B \times B') = \mathbf{M}_\varrho(x, B) \times \mathbf{M}_\varrho(x', B').$$

Returning to the proof of Theorem 4.1, note that if $Q_i := \{x'_{iN} < x_{iN}\}$, then $\{\tau_\varrho(x, x') > k \cdot N\} \subset \cap_{i=1}^k Q_i$. It is sufficient, therefore, to establish that

$$\mathbb{P} \cap_{i=1}^{j+1} Q_i \leq (1 - \varepsilon^2)\mathbb{P} \cap_{i=1}^{j} Q_i, \quad \forall j \in \mathbb{N}.$$

So pick any $j \in \mathbb{N}$. We have

$$\mathbb{P} \cap_{i=1}^{j+1} Q_i = \mathbb{P}(\mathbb{P}(\cap_{i=1}^{j+1} Q_i \,|\, \mathscr{F}_{jN})) = \mathbb{P}(\cap_{i=1}^{j} Q_i \mathbb{P}(Q_{j+1} \,|\, \mathscr{F}_{jN})).$$

We need only show that $\mathbb{P}(Q_{j+1} \,|\, \mathscr{F}_{jN}) \leq (1 - \varepsilon^2)$, or, equivalently, that

$\mathbb{P}(Q_{j+1}^c \,|\, \mathscr{F}_{jN}) \geq \varepsilon^2$. But

$$Q_{j+1}^c \supset \{x_{(j+1)N} \leq c\} \cap \{x'_{(j+1)N} \geq c\}$$

and by the Markov property we have

$$\mathbb{P}(\{x_{(j+1)N} \leq c\} \cap \{x'_{(j+1)N} \geq c\} \,|\, \mathscr{F}_{jN}) = \hat{\mathbf{M}}_\varrho^N(x_{jN}, x'_{jN}, [a, c] \times [c, b]),$$

which is

$$\mathbf{M}_\varrho^N(x_{jN}, [a, c]) \times \mathbf{M}_\varrho^N(x'_{jN}, [c, b]) \geq \mathbf{M}_\varrho^N(b, [a, c]) \times \mathbf{M}_\varrho^N(a, [c, b]),$$

which in turn is greater than or equal to ε^2 by Assumption 4.2.

5 Application: Optimal Growth

Recall the neoclassical infinite horizon economy of Brock and Mirman (1972). At time t income y_t is observed, a savings decision k_t is made, the current shock ξ_t is then revealed to the agent, and production takes place, realizing at the start of $t+1$ random output $y_{t+1} = f(k_t, \xi_t)$, which is net of depreciation. The process then repeats.

Preferences are specified by period utility function u and discount factor $\varrho \in (0, 1)$. Define Π to be the set of all feasible savings policies, which are Borel functions π from the set of positive reals to itself satisfying $\pi(y) \leq y$ for all y. Each $\pi \in \Pi$ determines a Markov process for income $(y_t)_{t=0}^\infty$ via (1), which in this case is

$$y_{t+1} = f(\pi(y_t), \xi_t) \tag{6}$$

where y_0 is given. The optimal investment problem is to solve (2), which is now

$$\max_{\pi \in \Pi} \mathbb{E} \sum_{t=0}^{\infty} \varrho^t u(y_t - \pi(y_t)),$$

with $(y_t)_{t=0}^\infty$ given by (6).

Assumption 5.1 *The utility function u is strictly increasing, differentiable, bounded and strictly concave, with $\lim_{c \to 0} u'(c) = \infty$.*

Assumption 5.2 *The sequence $(\xi_t)_{t=0}^{\infty}$ is IID on probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with cumulative distribution function G on $\mathbb{R}$. We suppose that there exists a $\underline{\xi} \in \mathbb{R}$ with $0 < G(x) < 1$ for all $x > \underline{\xi}$.*

Assumption 5.3 *The production function $f \colon [0, \infty) \times \mathbb{R} \to [0, \infty)$ satisfies the following assumptions. The map $k \mapsto f(k, z)$ is bounded, strictly increasing, strictly concave and continuously differentiable on $(0, \infty)$, with $f(0, z) = 0$ and $\lim_{k \to 0} f'(k, z) = \infty$, for each $z \in \mathbb{R}$. The map $z \mapsto f(k, z)$ is measurable and $\lim_{z \to \infty} f(k, z) = \infty$ for all $k \in (0, \infty)$.*

Most of the assumptions in 5.1–5.3 are standard. Requiring that the limit $\lim_{z \to \infty} f(k, z) = \infty$ for all $k \in (0, \infty)$ incorporates the most common case where the shock is multiplicative. Under this assumption the state space must be all of the positive reals.

In economics it is common to take the utility function as unbounded above (although rigorous justification of the dynamic programming arguments is not always provided). Note that if u has this property, then value boundedness never holds. The reason is that $v_\varrho(y) \geq u(y)$, so for fixed y' the difference

$$v_\varrho(y) - v_\varrho(y') \geq u(y) - v_\varrho(y')$$

cannot be bounded above by any constant.

We have also required that $k \mapsto f(k, z)$ is bounded. Whether or not this assumption can be relaxed is a more subtle issue. We leave it as an important open question.

We also need the following technical condition to manipulate the Euler equation. It holds in many situations we wish to consider (for example, when ξ_t is multiplicative and lognormally distributed).

Assumption 5.4 *Together, f and G satisfy $\int f'(k,z)^{-1}G(dz) < \infty$ for all $k \in (0,\infty)$.*

It is well-known that under our assumptions there is a unique DR-ϱ-optimal policy $\pi_\varrho \in \Pi$ for each $\varrho \in (0,1)$. Moreover, the DR-ϱ-optimal policy π_ϱ is pointwise increasing in ϱ (Danthine and Donaldson, 1981, Theorem 5.1). In other words, agents who discount the future more slowly invest more in all states. Given this monotonicity, we can always define $\pi_1 := \lim_{\varrho\uparrow 1}\pi_\varrho$. It is natural to conjecture that π_1 is AR-optimal.[11] In this connection, Dutta's result (1991, Theorem 3) shows that for this to be the case it is sufficient that the program satisfies value boundedness.

The main result of this section is

Proposition 5.1 *Under Assumptions 5.1–5.4, the stochastic neoclassical growth model is value bounded.*

For the proof we wish to apply Corollary 3.1. Fix $\hat{\varrho} \in (0,1)$, and let $\varrho \in [\hat{\varrho}, 1)$. Let S be the space $(0,\infty)$, and let x and x' be any two initial conditions. Let π_ϱ be the unique DR-ϱ-optimal policy associated with ϱ. Consider two economies with identical structure (u, f, G), both of which discount future utility according to ϱ and follow policy π_ϱ. The first has initial condition $y_0 = x$, and is perturbed by the sequence of shocks $(\xi_t)_{t=0}^\infty$, with $(y_t)_{t=0}^\infty$ defined by (6). The second has initial condition $y'_0 = x'$, and is perturbed by independent sequence $(\xi'_t)_{t=0}^\infty$, with $(y'_t)_{t=0}^\infty$ defined by (6).

The value function v_ϱ is known to be increasing (c.f., e.g., Mirman and Zilcha, 1975). As a result, if $x \leq x'$ then $\tau_\varrho(x,x') \equiv 0$. Suppose instead that $x' < x$.

[11] The interpretation of AR-optimality for the optimal growth model is clearest when $(y_t)_{t=0}^\infty$ is ergodic. In that case the sequence $\mathbb{E}\, u(y_s - \pi \circ y_s)$ and then the average $\left[\frac{1}{t}\sum_{s=0}^{t-1}\mathbb{E}\, u(y_s - \pi \circ y_s)\right]$ converge to the integral $\int u(y - \pi(y))F^*_\pi(dy)$, where F^*_π is the ergodic distribution corresponding to π. Then AR-optimality becomes equivalent to maximizing expected utility of consumption at the stochastic steady state—a generalization of the Phelps–Solow golden rule.

Although for our model there is a positive probability that y'_t exceeds y_t in every period, that probability will be arbitrarily small if y'_{t-1} is much smaller than y_{t-1}. However, we will show that $(y'_t)_{t=0}^{\infty}$ must return to a set (c, ∞) infinitely often, where $c > 0$, and once in that set there is an $\varepsilon > 0$ probability that y'_t exceeds y_t in the following period. As a result, we show that $\mathbb{P}\{\tau_\varrho(x, x') > t\} \to 0$ at a geometric rate depending only on x', and this is sufficient for Corollary 3.1. The details follow.

The first step concerns construction of the set (c, ∞) with the properties discussed above. We do this using a "Lyapunov" technique.

Lemma 5.1 *There are positive constants λ, β and a decreasing, real valued function w on $(0, \infty)$, all independent of ϱ, x and x', such that* (i) $w \geq 1$, (ii) $w(y) \to \infty$ *as* $y \to 0$, (iii) $\lambda < 1$, *and*

$$\mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \leq \lambda \cdot w \circ y'_t + \beta \quad \mathbb{P}\text{-}a.s. \tag{7}$$

Before beginning the proof, note that from (6) it is intuitively clear (and follows formally from the Markov property) that if w is any bounded or nonnegative real function then our time t prediction of the value $w \circ y_{t+1}$ satisfies

$$\mathbb{E}[w \circ y_{t+1} \,|\, \mathscr{F}_t] = \int w[f(\pi_\varrho(y_t), z)] G(dz) \quad \mathbb{P}\text{-a.s.} \tag{8}$$

A similar relation holds for y'_t and y'_{t+1}.

Also, we have an Euler equation to work with:

Theorem 5.1 (Mirman and Zilcha (1975)) *For each $\varrho \in (0, 1)$, the value function v_ϱ is concave and differentiable, and the optimal policy π_ϱ is interior. Let $y \in S$ and $\varrho \in (0, 1)$. For $c_\varrho(y) := y - \pi_\varrho(y)$ we have*

$$u'(c_\varrho(y)) = \varrho \int u'(c_\varrho(f(\pi_\varrho(y), z))) f'(\pi(y), z) G(dz). \tag{9}$$

Both π_ϱ and c_ϱ are increasing functions of y.

PROOF. [Proof of Lemma 5.1] We use an argument which draws on Nishimura and Stachurski (2005, Proposition 4.2). Our first candidate for w is the map $w(y) = \sqrt{u' \circ c_{\hat{\varrho}}(y)}$. By the Cauchy-Schwartz inequality we have

$$\int \sqrt{u' \circ c_{\hat{\varrho}}(f(\pi_{\hat{\varrho}}(y), z)) f'(\pi_{\hat{\varrho}}(y), z)} \sqrt{\frac{1}{f'(\pi_{\hat{\varrho}}(y), z)}} G(dz)$$
$$\leq \sqrt{\int u' \circ c_{\hat{\varrho}}(f(\pi_{\hat{\varrho}}(y), z)) f'(\pi_{\hat{\varrho}}(y), z) G(dz) \times \int \frac{1}{f'(\pi_{\hat{\varrho}}(y), z)} G(dz)}.$$

Using the definition of w and the Euler equation gives

$$\int w(f(\pi_{\hat{\varrho}}(y), z)) G(dz) \leq w(y) \sqrt{\int \frac{1}{\hat{\varrho} f'(\pi_{\hat{\varrho}}(y), z)} G(dz)}.$$

Using our assumptions and the Dominated Convergence Theorem, one can show that given $\lambda \in (0, 1)$, there exists a $\delta > 0$ such that

$$y < \delta \implies \sqrt{\int \frac{1}{\hat{\varrho} f'(\pi_{\hat{\varrho}}(y), z)} G(dz)} \leq \lambda.$$

Therefore,

$$y < \delta \implies \int w(f(\pi_{\hat{\varrho}}(y), z)) G(dz) \leq \lambda w(y).$$

Since $\varrho \geq \hat{\varrho}$ implies $\pi_{\varrho}(y) \geq \pi_{\hat{\varrho}}(y)$ for all y, we can in fact say that

$$y < \delta \implies \int w(f(\pi_{\varrho}(y), z)) G(dz) \leq \lambda w(y), \quad \forall \varrho \geq \hat{\varrho}.$$

In addition,

$$y \geq \delta \implies \int w(f(\pi_{\varrho}(y), z)) G(dz) \leq \int w(f(\pi_{\hat{\varrho}}(y), z)) G(dz)$$
$$\leq \int w(f(\pi_{\hat{\varrho}}(\delta), z)) G(dz) =: \beta < \infty.$$

Putting it together we get

$$\int w(f(\pi_{\varrho}(y), z)) G(dz) \leq \lambda w(y) + \beta, \quad \forall y \in S, \ \forall \varrho \geq \hat{\varrho}. \tag{10}$$

This in turn implies that $\forall \varrho \geq \hat{\varrho}$ we have

$$\int w(f(\pi_{\varrho}(y_t'), z)) G(dz) \leq \lambda w(y_t') + \beta. \tag{11}$$

Using the relation (8) we finish with

$$\mathbb{E}[\, w \circ y'_{t+1} \,|\, \mathscr{F}_t] \leq \lambda w \circ y'_t + \beta \quad \mathbb{P}\text{-a.s.} \tag{12}$$

as was to be shown. Note that w, β and λ are all independent of ϱ.

The only claim of Lemma 5.1 which is still in doubt is that $w \geq 1$. For $w(y) := \sqrt{u' \circ c_{\hat{\varrho}}}$ this is not necessarily true. However, we can replace w with $\hat{w} := w + 1$ if necessary, because when the bound (7) holds for w, λ and β then it also holds for $\hat{w}$, $\hat{\lambda} := \lambda$ and $\hat{\beta} := \beta + 1 - \lambda$. To see this, observe that

$$\begin{aligned}
\mathbb{E}[\, \hat{w} \circ y'_{t+1} \,|\, \mathscr{F}_t] &= \mathbb{E}[\, w \circ y'_{t+1} \,|\, \mathscr{F}_t] + 1 \\
&\leq \lambda w \circ y'_t + \beta + 1 \\
&= \lambda (w \circ y'_t + 1) + \beta + 1 - \lambda =: \hat{\lambda}\, \hat{w} \circ y'_t + \hat{\beta}.
\end{aligned}$$

All claims in the Lemma have now been verified.

The following corollary is an easy consequence of Lemma 5.1. From it we can infer that $(y'_t)_{t=0}^{\infty}$ must return relatively quickly to (c, ∞).

Corollary 5.1 *There is a constant $c > 0$ and an $\alpha \in (0, 1)$, both independent of ϱ, x and x', such that*

$$\mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \cdot \mathbb{1}\{y'_t \leq c\} \leq \alpha \cdot w \circ y'_t \cdot \mathbb{1}\{y'_t \leq c\}. \tag{13}$$

PROOF. By (ii) there is a $c > 0$ such that $w(c) > \beta(1 - \lambda)^{-1}$. Since w is decreasing, $w(x) \geq w(c)$ for all $x \in (0, c]$. Define

$$\alpha := \lambda + \frac{\beta}{w(c)},$$

so that $\lambda < \alpha < 1$. By Lemma 5.1, then,

$$\mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \cdot \mathbb{1}\{y'_t \leq c\} \leq (\lambda \cdot w \circ y'_t + \beta) \cdot \mathbb{1}\{y'_t \leq c\}.$$

$$\therefore \quad \frac{\mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \cdot \mathbb{1}\{y'_t \leq c\}}{w \circ y'_t} \leq \left(\lambda + \frac{\beta}{w \circ y'_t}\right) \mathbb{1}\{y'_t \leq c\}$$
$$\leq \alpha \mathbb{1}\{y'_t \leq c\}.$$

Let $N_t := \sum_{i=0}^{t} \mathbb{1}\{y'_i > c\}$, so that N_t is the number of times $y'_i > c$ in the period $0, \ldots, t$. Fix $j \leq t$. Omitting the subscript ϱ, we have

$$\mathbb{P}\{\tau > t\} = \mathbb{P}\{\tau > t\} \cap \{N_t > j\} + \mathbb{P}\{\tau > t\} \cap \{N_t \leq j\}. \tag{14}$$

The two terms on the right hand side need to be bounded.

It is convenient to begin with the second term in (14). For this purpose, let $B := \alpha^{-1} \int w[f(\pi_{\hat{\varrho}}(c), z)]G(dz)$, which can be shown to be finite using (7). Next, let $M_t := \alpha^{-t} B^{-N_{t-1}} \, w \circ y'_t$, where $N_{-1} := 0$, so $M_0 = w \circ y'_0 \equiv w(x')$.

Lemma 5.2 *The sequence $(M_t)_{t=0}^{\infty}$ is a supermartingale with respect to the filtration $(\mathscr{F}_t)_{t=0}^{\infty}$.*

PROOF. Clearly M_t is $\mathscr{F}_t$-measurable. It will be integrable provided that we can verify the key supermartingale property $\mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \leq M_t$. To this end, let $F := \mathbb{1}\{y'_t > c\}$ and $F^c := 1 - F = \mathbb{1}\{y'_t \leq c\}$, so that

$$\mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] = \mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F + \mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F^c.$$

Consider the first term. On F we have $N_t = N_{t-1} + 1$, so

$$\begin{aligned}
\mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F &= \alpha^{-(t+1)} B^{-N_{t-1}} B^{-1} \mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \cdot F \\
&= \alpha^{-(t+1)} B^{-N_{t-1}} B^{-1} \int w(f(\pi_{\varrho} \circ y'_t, z)) G(dz) \cdot F \\
&\leq \alpha^{-(t+1)} B^{-N_{t-1}} B^{-1} \int w(f(\pi_{\hat{\varrho}}(c), z)) G(dz) \cdot F \\
&\leq \alpha^{-t} B^{-N_{t-1}} F.
\end{aligned}$$

Using this bound and $w \geq 1$ gives $\mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F \leq M_t \cdot F$. Also, on the set

F^c we have $N_t = N_{t-1}$, and Corollary 5.1 applies. Hence,

$$\begin{aligned}\mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F^c &= \alpha^{-t} B^{-N_{t-1}} \alpha^{-1} \mathbb{E}[w \circ y'_{t+1} \,|\, \mathscr{F}_t] \cdot F^c \\ &\leq \alpha^{-t} B^{-N_{t-1}} w \circ y'_t \cdot F^c.\end{aligned}$$

$$\therefore \quad \mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \cdot F^c \leq M_t \cdot F^c.$$

$$\therefore \quad \mathbb{E}[M_{t+1} \,|\, \mathscr{F}_t] \leq M_t.$$

In view of the supermartingale property we have $\mathbb{E} M_t \leq \mathbb{E}\, M_0 = w(x')$, whence

$$\begin{aligned}\mathbb{P}\{\tau > t\} \cap \{N_t \leq j\} &\leq \mathbb{P}\{N_{t-1} \leq j\} \\ &= \mathbb{P}\{B^{-N_{t-1}} \geq B^{-j}\} \quad (\because B \geq 1) \\ &\leq B^j \, \mathbb{E}\, B^{-N_{t-1}} \quad (\because \text{ Chebychev's ineq.}) \\ &\leq \alpha^t B^j \, \mathbb{E}\, M_t \qquad (\because \;\; w \geq 1) \\ &\leq \alpha^t B^j \, w(x').\end{aligned}$$

Now we return to the first term in (14), which has the following simple bound.

Lemma 5.3 *There is an $\varepsilon > 0$ independent of ϱ, x and x' such that*

$$\mathbb{P}\{\tau > t\} \cap \{N_t > j\} \leq (1 - \varepsilon)^j.$$

The intuition is that whenever $y'_t > c$ the income ranking reverses with independent probability at least ε. Before starting on the proof, let σ_j be the time of the j-th visit of $(y'_t)_{t=0}^{\infty}$ to (c, ∞). We can define these random variables recursively by $\sigma_1 := \inf\{t \geq 0 : y'_t > c\}$, and

$$\sigma_{j+1} := \inf\{t \geq \sigma_j + 1 : y'_t > c\}.$$

All of these stopping times are finite $\mathbb{P}$-a.s., because $\{\sigma_j = \infty\} \subset \cap_{t=0}^{\infty} \{N_t < j\}$, and $\mathbb{P}\{N_t < j\} \leq \alpha^t B^{j+1} w(x')$, as was previously shown.

PROOF. [Proof of Lemma 5.3] Let $Q_i := \{y'_{\sigma_i + 1} < y_{\sigma_i + 1}\}$. In other words, Q_i is the event that no swap occured in the period after the i-th visit of $(y'_t)_{t=0}^{\infty}$

to (c, ∞). For Q_i so defined, we have

$$\mathbb{P}\{\tau > t\} \cap \{N_t > j\} \leq \mathbb{P} \cap_{i=1}^{j} Q_i. \tag{15}$$

To see this, observe that when $N_t > j$ we have $\sigma_j < t$. If $\tau > t$ is also true, we know that neither this visit to (c, ∞) nor any of the previous ones resulted a reversal of incomes. In other words, the statement $\cap_{i=1}^{j} Q_i$ is true.

It therefore suffices to bound $\mathbb{P} \cap_{i=1}^{j} Q_i$. To this end, let

$$W := \{(z, z') \in \mathbb{R}^2 : f(\pi_{\hat{\varrho}}(c), z') \geq \lim_{x \to \infty} f(x, z)\}.$$

It is not difficult to see that $Q_i \subset \{(\xi_{\sigma_i}, \xi'_{\sigma_i}) \notin W\}$, because if $(\xi_{\sigma_i}, \xi'_{\sigma_i}) \in W$, then $y'_{\sigma_i+1} = f(\pi_\varrho(y'_{\sigma_i}), \xi'_{\sigma_i}) \geq f(\pi_{\hat{\varrho}}(y'_{\sigma_i}), \xi'_{\sigma_i}) \geq f(\pi_{\hat{\varrho}}(c), \xi'_{\sigma_i}) \geq \lim_{x \to \infty} f(x, \xi_t) \geq f(\pi_\varrho(y_{\sigma_i}), \xi_{\sigma_i}) = y_{\sigma_i+1}$. It is also clear that $\varepsilon := \mathbb{P}\{(\xi_t, \xi'_t) \in W\}$ is a strictly positive number independent of ϱ.[12]

Now suppose we can show for this ε that $\mathbb{P}(\cap_{i=1}^{k+1} Q_i) \leq (1 - \varepsilon)\mathbb{P}(\cap_{i=1}^{k} Q_i)$ holds for any k. In view of (15) this will complete the proof, as iterating backwards gives $\mathbb{P} \cap_{i=1}^{j} Q_i \leq (1 - \varepsilon)^j$.

So pick any $k \in \mathbb{N}$. We have

$$\mathbb{P} \cap_{i=1}^{k+1} Q_i = \mathbb{P}(\mathbb{P}(\cap_{i=1}^{k+1} Q_i \,|\, \mathscr{F}_{\sigma_k+1})) = \mathbb{P}(\cap_{i=1}^{k} Q_i \mathbb{P}(Q_{k+1} \,|\, \mathscr{F}_{\sigma_k+1})).$$

We need only show that $\mathbb{P}(Q_{k+1} \,|\, \mathscr{F}_{\sigma_k+1}) \leq (1 - \varepsilon)$, or, equivalently, that $\mathbb{P}(Q^c_{k+1} \,|\, \mathscr{F}_{\sigma_k+1}) \geq \varepsilon$. But $Q^c_{k+1} \supset \{(\xi_{\sigma_{k+1}}, \xi'_{\sigma_{k+1}}) \in W\}$, which is independent of $\mathscr{F}_{\sigma_k+1}$ and has probability ε. The proof is done.

Let's now complete the proof of Proposition 5.1. Choose $n \in \mathbb{N}$ such that

[12] To verify strict positivity, choose $N \in \mathbb{N}$ s.t. $\nu\{z \in \mathbb{R} : b(z) \leq N\} > 0$, where $b(z) := \lim_{x \to \infty} f(x, z)$. Since $(\xi_t, \xi'_t) \in W$ is implied by $f(\pi_{\hat{\varrho}}(c), \xi'_t) \geq N$ and $b(\xi_t) \leq N$, and since these last two events are independent and have strictly positive probability, it follows that $\varepsilon > 0$.

$\delta := \alpha^n B < 1$, and set $j = t/n$, so that $\alpha^t B^j = \delta^{t/n}$.

$$\begin{aligned}
\mathbb{E}\tau_\varrho(x,x') &\le \sum_{t=0}^{\infty} t\mathbb{P}\{\tau_\varrho(x,x') \ge t\} \\
&= \sum_{t=0}^{\infty} (t+1)\mathbb{P}\{\tau_\varrho(x,x') > t\} \\
&\le \sum_{t=1}^{\infty} (t+1)[(1-\varepsilon)^{t/n} + \delta^{t/n} w(x')].
\end{aligned}$$

$$\therefore \quad \mathbb{E}\tau_\varrho(x,x') \le S + T \cdot w(x'),$$

where constants S and T are independent of ϱ and x. The conditions of Corollary 3.1 are therefore met.

Appendix

Here we justify $\mathbb{E}\sum_{t=\tau}^{\infty} \varrho^t r(x_t, \pi(x_t)) = \mathbb{E}\,\varrho^\tau v(x_\tau)$ from the proof of Theorem 3.1. We assume the reader is familiar with stopping times, the strong Markov property and composition of Markov kernels. We set $\mathbf{M}^t g(x)$ to be $\int g(y)\mathbf{M}^t(x,dy)$. For all of the underlying theory and other notation see, for example, Durrett (1996).

We can and do assume that $\tau_\varrho(x_0, x_0')$ is finite $\mathbb{P}$-almost surely (otherwise the bound in Theorem 3.1 is trivial). We then have

$$\begin{aligned}
\mathbb{E}\sum_{t=\tau}^{\infty} \varrho^t r(x_t,\pi(x_t)) &= \mathbb{E}\left[\mathbb{E}\left[\sum_{t=\tau}^{\infty} \varrho^t r(x_t,\pi(x_t)) \,|\, \mathscr{F}_\tau\right]\right] \\
&= \mathbb{E}\left[\sum_{t=\tau}^{\infty} \varrho^t \mathbb{E}[r(x_t,\pi(x_t)) \,|\, \mathscr{F}_\tau]\right] \\
&= \mathbb{E}\left[\varrho^\tau \sum_{t=0}^{\infty} \varrho^t \mathbf{M}^t r(x_\tau,\pi(x_\tau))\right] = \mathbb{E}\varrho^\tau v(x_\tau).
\end{aligned}$$

References

[1] Brock, W.A., Mirman, L.J., 1972. Optimal economic growth and uncertainty: the discounted case. Journal of Economic Theory 4, 479–513.

[2] Brock, W.A., Mirman, L.J., 1973. Uncertainty and economic growth: the no-discounting case. International Economic Review 14 (3), 560–573.

[3] Chiarella, C., Szidarovsky, F., 2001. The nonlinear Cournot model under uncertainty with continuously distributed time lags. Central European Journal of Operations Research 9 (3), 183–196.

[4] Danthine, J-P., Donaldson, J.B., 1981. Stochastic properties of fast vs. slow growth economies. Econometrica 49 (4), 1007–1033.

[5] Durrett, R., 1996. Probability: Theory and Examples. 2nd ed., Duxbury Press, California.

[6] Dutta, P.K., 1991. What do discounted optima converge to? A theory of discount rate asymptotics in economic models. Journal of Economic Theory 55, 64–94.

[7] Hopenhayn, H.A., Prescott, E.C., 1992. Stochastic monotonicity and stationary distributions for dynamic economies. Econometrica 60, 1387–1406.

[8] Jaśkiewicz A., Nowak, A.S., 2006. On the optimality equation for average cost Markov control processes with Feller transition probabilities. Journal of Mathematical Analysis and Applications 316, 495–509.

[9] McKenzie, L.W., 1998. Turnpikes. American Economic Review 88 (2), 1–14.

[10] Mirman, L.J., Morand, O.F. and Reffett, K. 2005. A qualitative approach to Markovian equilibrium in infinite horizon economies with capital, manuscript.

[11] Mirman, L.J., Zilcha, I., 1975. On optimal growth under uncertainty. Journal of Economic Theory 11, 329–339.

[12] Nishimura, K., Stachurski, J., 2005. Stability of stochastic optimal growth models: A new approach. Journal of Economic Theory 122 (1), 100–118.

[13] Ramsey, F., 1928. A mathematical theory of saving. Economic Journal 38, 543–559.

[14] Rosenthal, J.S., 2003. Asymptotic variance and convergence rates of nearly-periodic MCMC algorithms. Journal of the American Statistical Association 98, 169–177.

[15] Sennott, L.I., 1986. A new condition for the existence of optimal stationary policies in average cost Markov decision processes. Operations Research Letters 5, 17–23.

[16] Stokey, N.L., Lucas, R.E., and Prescott, E.C., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Massachusetts.