Journal of Economic Behavior & Organization 79 (2011) 194–210

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Classical turnpike theory and the economics of forestry☆

M. Ali Khan[a,b,*], Adriana Piazza[c]

[a] Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, United States
[b] Department of Economics, University of Queensland, Brisbane, QLD 4072, Australia
[c] Departamento de Matemática, Universidad Técnica Federico Santa María. Avda. España 1680, Casilla 110-V, Valparaíso, Chile

ARTICLE INFO

Article history:
Received 18 August 2009
Received in revised form 28 January 2011
Accepted 31 January 2011
Available online 3 March 2011

JEL classification:
C62
D90
Q23

Keywords:
Turnpike theory
Asymptotic convergence
Good program
Approximately optimal program
Time horizon
Large but finite
Non-interiority
Nondifferentiability
Forest management

ABSTRACT

Classical turnpike theory, as originally conceived by Samuelson, pertains to optimal growth theory over a large but finite time horizon with given initial and terminal stocks. In this paper, we present two turnpike results in the context of the economics of forestry with given initial and terminal forest configurations. Our results depart from the general theory in that they concern a transitional production set which does not satisfy the assumptions of *inaction* and *free disposal*, and rely on a recently discovered non-interiority assumption on concave (not necessarily differentiable) benefit functions that implies, and is implied by, the asymptotic convergence of good programs.

The [turnpike] results while simple and concise could probably not be predicted in advance. (Gale 1970)

1. Introduction

The results presented in this paper can perhaps be best introduced by asking what is “classical turnpike theory”? and what has it to do with the “economics of forestry”? We begin with the first question.

☆ This work was initiated during Piazza’s visit to Johns Hopkins in May 2008, continued during her visit to the University of Illinois at Urbana-Champaign in April–May, 2009, and completed when Khan held the position of Visiting Professor at the University of Queensland, June–July 2009, and at the Nanyang Technological University, August, 2009. The authors thank Joseph Harrington, Flavio Menezes, Euston Quah, Anne Villamil and Nicholas Yannelis, and each department, for their hospitality. They also thank Tapan Mitra for correspondence and conversation, and two anonymous referees of this journal for their suggestions. Adriana Piazza gratefully acknowledges the financial support of FONDECYT under Grant # 3080059 and Grant # 11090254 and that of Programa Basal PFB 03, CMM, U. de Chile.

* Corresponding author at: Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, United States. Tel.: +1 410 516 8545; fax: +1 410 516 7600.
E-mail addresses: akhan@jhu.edu (M.A. Khan), adriana.piazza@usm.cl (A. Piazza).

0167-2681/$ – see front matter
doi:10.1016/j.jebo.2011.01.030

The origin of the subject is easily dated to Paul Samuelson's (1949) Rand Memorandum, and its essential motivation is to show that the solutions of an intertemporal allocation problem over a large but finite time horizon, with given initial and terminal stocks, stay approximately close, most of the time, to maximally sustainable programs free from any considerations pertaining to a terminal horizon. However, despite the commitment in Ramsey (1928), it is only by 1965 that an infinite-horizon variational problem is accepted as a legitimate vehicle to address issues of intertemporal allocation. In an undiscounted setting, Gale (1967) focussed on the problem of the existence of an optimal program, and resolved it by conditions ensuring the asymptotic convergence of *good* programs, a methodological procedure that he referred to as a "round-about method."[1] Even though this asymptotic property of good programs is mentioned in passing as a "turn-pike" theorem, Gale reverted to the classical conception in his 1970 expository paper. Along with monotonicity properties, he situates classical turnpike theory within the broader rubric of qualitative properties of optimal programs.[2] In this paper, we revert to the Samuelsonian classical conception pertaining to large but finite optimal programs, and investigate what McKenzie (1976) was subsequently to term the *middle* turnpike.[3]

The resurgence of interest in the economics of forestry can also be dated to Samuelson (1976a) and to his sighting of Faustman's (1849) analysis. However, it remained for the remarkable articles of Mitra and Wan (1985, 1986), and following them, Salo and Tahvonen (2002, 2003), to take his market analysis and recast it into an optimal Ramseyian planning, framework. Once this link is forged, and the conceptual markers of the theory of intertemporal resource allocation are identified in the forestry model – the golden-rule stocks, the golden-rule prices and corresponding maximal sustainable timber yields as Ramsey's bliss point – we are naturally lead to ask whether forest management programs for large but finite time horizons follow a turnpike. To go to the epigraph from Gale, there appears to be no reason why they should, and this scepticism leads directly into our second question concerning the relevance of classical turnpike theory to the Mitra–Wan (MW) forestry model. Here the usual justifications of turnpike theory as resolving qualitative puzzles and computational difficulties take on an added force, and turnpike theory attains a normative significance perhaps even greater than that in capital theory in the abstract and in the large. And to be sure, the resulting theory includes, as a special case, the situation of a profit-maximizing forest manager, with the infinite horizon as the formalization of the fact that there is no specified time-period for the end of his or her firm, with such an application involving a demand function for timber being very much facilitated under the weakened assumptions on benefit (felicity) functions that we pursue in this paper.[4]

Unlike the aggregative Ramsey–Cass–Koopmans optimal growth model of modern macroeconomics, the issue in the economics of forestry is the lack of knowledge of the optimal policy correspondence and of transition dynamics. What should the planner, or the forest manager, do today in terms of optimal policy over a fixed time-horizon rather than take satisfaction that any arbitrary initial configuration converges to the maximally sustainable forest configuration in the very long run? In any case, at least since Faustman, periodicity seems to be the rule, and the charting of an optimal policy correspondence with a non-linear benefit (felicity) function remains, even until now, totally uncharted.[5] And so the question arises whether staying with the maximally sustainable configuration is "good enough," and thereby leads only to an increased reliance on turnpike results.[6] Furthermore, a forest, unlike a given stock of capital, is much more than a durable input for the production of desired commodities – it is desirable in itself, and if not a "way of life" of so-called endogenous and native communities, a stock imbricated by externality considerations and entrusted by one generation to another.[7] As such, a forest configuration also enters as an argument in the benefit (felicity) function, and thereby further complicates the difficulties of determining what the planner has to do "tomorrow and the day after" rather than the long-run. But more to the point, such results allay fears and furnish a reassurance that a planner's departure from an initial forest configuration to one yielding maximal sustained timber yields, even when he has to return to future generations the forest in the same state that he was given it, is not betraying this trust. However such a theory can only be constructed on the shoulders of a theory for the simpler case considered here when the felicities depend only on timber yields.[8]

With the relevance of turnpike theory to the economics of forestry established and out of the way, one can turn to the more immediately antecedent literature and delineate the precise contribution of this paper. Given the extensive work on turnpike theory associated with the names of Samuelson, McKenzie, Gale and their followers, why can one not simply appeal

[1] Gale (1967, p.1) writes, "It may well be that there is a more direct way of obtaining our existence theorem, but even if this should turn out to be true, the present round-about approach would not be an entirely wasted effort." The continuing relevance of this "round-about" method is noted, and utilized, in Khan and Zaslavski (2010a,b).

[2] See paragraph 3 and Section 5 in Gale (1970). (The reader is warned that there are two Sections 4 in the paper.)

[3] For an argument that seeks to distinguish the classical conception from the broader one of McKenzie that emphasizes both the *early* and *late* turnpike, see Khan and Zaslavski (2010a). McKenzie's followers now refer to asymptotic convergence of the late turnpike exclusively as *the* turnpike theory.

[4] Such an application was already noted in Mitra and Wan (1986). For a detailed consideration of alternative objectives of the forest manager, the reader is referred to Samuelson (1976a) and Rosser (2005).

[5] Salo and Tahvonen (2002) and Mitra (2006) are notable exceptions for the strictly concave case, but even they confine themselves to a dual-aged forest and require additional assumptions. For a result on a periodic turnpike, see Samuelson (1976b).

[6] Also see the last paragraph of Gale's (1967) introduction. Its somewhat "paradoxical" defense of infinite-horizon problems as pertaining to the very immediate future – the "guidance of a ship on a long journey" – is premised on precisely the ability to compute this policy correspondence.

[7] For externality considerations in the economics of forestry, see Samuelson (1976a); and for the larger implications reaching into political theory, see Kant and Berry (2005), and perhaps also the discussion of the references in Khan (2005). In a non-Ramseyian framework, the complications arising from externalities are discussed in Rosser (2005).

[8] For the complications arising from the case of wealth-dependent felicities, see Majumdar and Mitra (1994).

to the standard results? What is the need for additional modifications? In particular, why are the results presented in this paper not straightforward applications of the recent extension of the theory sketched in Khan and Zaslavski (2010a)? The answers to these questions require a technically more focussed discussion. We turn to this.

Even though the Robinson–Solow–Srinivasan (RSS) model and the MW model are different models with entirely different interpretive registers, the subtle analytical connections between them are undeniable, and the recent RSS revisitation of turnpike theory in Khan and Zaslavski (2010b) worth noting. It involves at least four disparate elements: (i) the irrelevance, in principle, of necessary first order Euler–Lagrange conditions, and indeed of differentiability of the felicity function at the golden-rule stock, (ii) the identification of asymptotic convergence of good programs as a sufficient condition for classical turnpike theory, and therefore for the asymptotic convergence of optimal programs, (iii) the derivation of asymptotic stability of optimal programs from the classical turnpike result, which is to say, the derivation of results on the early and late turnpikes as a *consequence* of a result on the middle turnpike, (iv) a focus on approximately optimal large but finite programs. Only points (i) and (ii) need further supplementation in the context of results that we report here, and we take them in turn. As regards (i), it is now well-understood that the golden-rule stock in the RSS model, is not in the interior of the transition set, and even for the case with a single type of machine when it is in the interior, the reduced form utility function is not differentiable at it even with a linear felicity function. In part, this is precisely what gives the RSS model its continuing interest. The same occurs in the MW model. As regards (ii), it allows us to move away from the dichotomy of linear and strictly concave felicity functions to a more productive sufficiency condition, something essential for the RSS model where even strictly concave felicity functions do *not* lead to strictly concave reduced-form utility functions as is required by the theory. In particular, such a condition allows a turnpike theorem when the felicity function is linear and the marginal rate of transformation $\xi_\sigma \neq 1$. It is this issue that finds its most satisfactory culmination in the MW model.

In recent work, Khan and Piazza (2009) furnish for the MW forestry model a non-interiority condition that is necessary *and* sufficient for asymptotic convergence of good programs when the benefit (felicity) function is assumed only to be concave and not necessarily differentiable.[9] And so the natural question arises as to whether one can construct a robust turnpike theory of the classical type for the economics of forestry, as is done in Khan and Zaslavski (2010a) for the (RSS) choice of technique problem in development planning. Theorems A and B presented in Section 3 below answer this question in the affirmative, and constitute the principal results of this work. They also yield as direct corollaries results on the asymptotic convergence of optimal programs (Theorem 2.4), and results on the "bunching" or "approximate bunching" of optimal forest configurations. Unlike the RSS model, there is no natural ordering on the transition production set in the MW model, and this necessitates novel and different arguments, and as it happens, more constructive ones than those presented in Khan and Zaslavski (2010a). Sections 4 and 5 present the statements of several results that are both needed in the proofs of the results of Section 3, and are interesting in their own right. We take care to comment on aspects in which they depart from the corresponding arguments in Khan and Zaslavski (2010a); the formalities of the proofs themselves are confined to the Appendix A. Section 6 concludes the paper with a delineation of three directions and the open questions associated with each. In the next section, by way of introducing the reader to the notation and the terminology, we present the basic analytics of the model. This material is by now well-understood, but for the sake of completeness, we present results that being phrased in terms of our non-interiority condition, generalize corresponding results in Khan and Piazza (2010).

2. The Mitra–Wan tree farm and the non-interiority condition

We begin by introducing some notation. Let $\mathbb{N}$ be the set of non-negative integers and $\mathbb{R}$ ($\mathbb{R}_+$) the set of real (non-negative) numbers. We shall work in the $n-1$-dimensional simplex $\Delta = \{x \in \mathbb{R}_+ \sum_{i=1}^{n} x_i = 1\}$. For any $x, y \in \mathbb{R}^n$ we denote the inner product by $xy = \sum_{i=1}^{n} x_i y_i$ and the supreme norm of x by $||x||_\infty$.

In addition to its original formulation (Mitra and Wan, 1985, 1986), an outline of the Mitra–Wan forestry model is also available in Mitra (2005, 2006). Here we depart from the original specification and work with the reformulation presented in Salo and Tahvonen (2002) and pursued in Khan (2005), Khan and Piazza (2010) and Salo and Tahvonen (2003). Let us consider a forest of total area 1 occupied by trees of the same species. In contrast with the case of wild forests, the state of a forest plantation may be described by specifying the areas occupied by trees of different ages, the underlying assumption being that the timber content per unit of area is related only to the age of the trees.[10]

Under this specification, the model consists simply of the pair (b, w). The vector b is a non-negative vector of timber coefficients $(b_1, \ldots, b_n) \in \mathbb{R}^n_+$ where b_i represents the volume of timber contained in a unit of land occupied with trees of age i. Note that we make no assumptions on the timber coefficients other than the following Brock–Mitra–Wan uniqueness condition.

Standing Hypothesis (BMW): There exists $\sigma \in \{1, \ldots, n\}$ such that $(b_\sigma/\sigma > b_i/i)$ for all $i \in \{1, \ldots, n\}/\{\sigma\}$.

[9] After the results presented here were obtained, Tapan Mitra provided a transparent equivalent formulation of the necessary and sufficient condition presented as the non-interiority Condition 2.1 below in terms of the "concavity at a point" property.

[10] This assumption is a consequence of considering the growth of the trees as a pure aging process and that trees are sown within a constant distance from each other throughout the area.

The felicity function $w : [0, \infty) \to \mathbb{R}$ represents the instantaneous economic benefit as a function of the total volume of timber harvested at each time step. A forest configuration is an element of Δ, representing the fact that trees of ages ranging from one to n cover completely a homogeneous plot of land of normalized unit size.

In addition to this, we very much follow the original conception and assume that there are no costs of plantation, and that n is the age after which a tree dies or losses its economic value. However, one difference should be noted. In their treatment, Mitra–Wan take N to be the age at which the biomass per unit of land is maximized, claiming that “for any reasonable objective function for the economy, trees will never be allowed to grow beyond age N; we therefore take this as a condition of feasibility itself.”[11] It is this reasoning that allows the authors to limit themselves to an N-dimensional state vector. However, given the fact that a concave felicity function favors a homogeneously configured forest, the planner may well adopt the trade-off of postponing harvesting beyond age N in order to reshape the forest into a more homogeneous state. We circumvent this by assuming n to be the age at which a tree dies, and point out that the technicalities of the analysis do not change with this augmentation of the state's dimension.

For each period $t \in \mathbb{N}$ we denote $x_i(t) \geq 0$, $i = 1, \ldots, n$, the surface occupied by trees of age i at time t. We represent the state of the forest by the vector $x(t) = (x_1(t), \ldots, x_n(t)) \in \Delta$.

At every stage we must decide how much land to harvest of every age-class, $c(t) = (c_1(t), \ldots, c_n(t))$ where $c_i(t) \in [0, x_i(t)]$. As we know that after n a tree has no value, we assume that $c_n(t) = x_n(t)$ for all t. By the end of period $t+1$, the state will be exactly

$$x(t+1) = \left(\sum_{i=1}^{n} c_i(t), x_1(t) - c_1(t), \ldots, x_{n-1}(t) - c_{n-1}(t) \right).$$

Definition 2.1. A sequence $\{x(t)\}_{t=0}^{\infty}$ is called a program if for each $t \geq 0$

$$\begin{cases} x(t) \in \Delta, \\ x_{i+1}(t+1) \leq x_i(t) \quad i = 1, \ldots, n-1 \end{cases} \tag{1}$$

Definition 2.2. Let T_1 and T_2 be integers such that $0 \leq T_1 < T_2$. A sequence $\{x(t)\}_{t=T_1}^{t=T_2}$ is called a program if $x(T_2) \in \Delta$ and relations (1) hold for each t satisfying $T_1 \leq t < T_2$.

Define the transition possibility set Ω as the collection of pairs $(x, x') \in \Delta \times \Delta$ such that it is possible to go from the state x in the current period (today) to the state of the forest x' in the next period (tomorrow) fulfilling relations (1). Formally,

$$\Omega = \{(x, x') \in \Delta \times \Delta / x_i \geq x'_{i+1} \quad \text{for all}\, i = 1, \ldots, n-1\}$$

Definition 2.3. The vector of harvests needed to perform this transition is given by the function $\lambda : \Omega \to \mathbb{R}^n_+$,

$$\lambda(x, x') = (x_1 - x'_2, x_2 - x'_3, \ldots, x_{n-1} - x'_n, x_n)$$

In addition, it is easy to see that

$$(x, x') \in \Omega \Leftrightarrow x, x' \in \Delta \text{ and } \lambda(x, x') \geq 0$$

The preferences of the planner are represented by a felicity function, $w : [0, \infty) \to \mathbb{R}$ which is assumed to be continuous, strictly increasing and concave. Define for any $(x, x') \in \Omega$ the function $u(x, x')$ as

$$u(x, x') = w(bc) \,\text{where}\, c = \lambda(x, x')$$

Definition 2.4. A *golden-rule stock* $\hat{x} \in \mathbb{R}^n_+$ is such that $(\hat{x}, \hat{x})$ is a solution to the problem:

$$\begin{cases} \text{maximize} & u(x, x) \\ \text{subject to} & (x, x) \in \Omega \end{cases}$$

We now present some basic antecedent results, and except those indicated at the end of the section, they are all taken from Khan and Piazza (2009).

Theorem 2.1. *There exists a unique golden-rule stock* $\hat{x} = \left(\underbrace{\frac{1}{\sigma}, \ldots, \frac{1}{\sigma}}_{\sigma}, 0, \ldots, 0 \right)$

[11] See Mitra and Wan (1986, p. 232). The same point is made in Mitra (2005, Section 4, Paragraph 5).

We denote by $\hat{c}$ the vector of harvests obtained by the pair $(\hat{x}, \hat{x})$, namely $\hat{c} = \lambda(\hat{x}, \hat{x})$. Observe that $\hat{c}_\sigma = 1/\sigma$ and $\hat{c}_i = 0$ for all $i \neq \sigma$. The total amount of timber obtained by such a vector of harvest is $b\hat{c} = b_\sigma/\sigma$. Pick any $z \in \partial^+ w\left(b_\sigma/\sigma\right)$,[12] and set $\hat{p} \in \mathbb{R}^n_+$, $\hat{p} = zb_\sigma/\sigma(1, 2, \ldots, n) > 0$.

Definition 2.5. We define the *value loss* associated with any $(x, x') \in \Omega$ to be

$$\delta(x, x') = w\left(\frac{b_\sigma}{\sigma}\right) - w(b\lambda(x, x')) - \hat{p}(x' - x).$$

It is easy to see that the function $\delta(\cdot, \cdot)$ is convex and the following lemma asserts that $\delta(x, x') \geq 0$ for any $(x, x') \in \Omega$.

Lemma 2.1. *For any $(x, x') \in \Omega$ we have*

$$\delta(x, x') \geq z\left[\sum_{i=1}^{n-1}\left(\frac{b_\sigma}{\sigma} - \frac{b_i}{i}\right) i\,(x_i - x'_{i+1}) + \left(\frac{b_\sigma}{\sigma} - \frac{b_n}{n}\right) nx_n\right] \geq 0 \tag{2}$$

We use the following notion of *good* and *bad* programs introduced by Gale (1967)

Definition 2.6. A program $\{x(t)\}$ is called *good* if there exists $M \in \mathbb{R}$ such that for all $T \geq 0$, $\sum_{t=0}^{T}[w(bc(t)) - w(b_\sigma/\sigma)] \geq M$, where $c(t) = \lambda(x(t), x(t+1))$. A program is *bad* if $\lim_{T\to\infty}\sum_{t=0}^{T}[w(bc(t)) - w(b_\sigma/\sigma)] = -\infty$.

The following general result of Gale applies to the MW model.

Proposition 2.1. *Programs are partitioned into good and bad programs. Furthermore,*

i. *$\{x(t)\}$ is good iff $\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) < \infty$.*
ii. *$\{x(t)\}$ is bad iff $\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) = \infty$.*

Let $x_0 \in \Delta$. Set $\mu(x_0) = \inf\{\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) : \{x(t)\}$ is a program from $x_0\}$.

It is possible to see that there exists at least one good program from every $x_0 \in \Delta$, which in turn implies that $\mu(x_0) < \infty$. The following result can now be established.

Proposition 2.2. *From any $x_0 \in \Delta$ there exist a good program $\{x(t)\}$ such that*

$$\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) = \mu(x_0). \tag{3}$$

The fact that every good program converges to the golden rule stock in the case that w is strictly concave was established in Mitra and Wan (1986, Lemma 6.4). Khan and Piazza (2009) provide a necessary and sufficient condition to assure the convergence of every good program to the golden rule stock for any concave utility function w that is not necessarily differentiable. We describe this characterization in the following terms.

Let the *discrepancy function* f be

$$f(\xi) = w\left(\frac{b_\sigma}{\sigma}\right) - w(b_\sigma\xi) + z(b_\sigma\xi - \frac{b_\sigma}{\sigma}). \tag{4}$$

We can appeal to standard results in Rockafellar (1970) to assert that the concavity of w implies $f(1/\sigma) = 0$, $f(\xi) \geq 0$ for all ξ and f attains its minimum in a closed interval S_f containing $(1/\sigma)$.

We now turn to the condition that serves as a basic standing hypothesis for our principal results.

Condition 2.1 (Non-interiority). $(1/\sigma) \notin \operatorname{int} S_f$.

Either w coincides with the support function, $w(b_\sigma/\sigma) + z(\cdot - b_\sigma/\sigma)$ only at the point b_σ/σ or this point is one of the extremes of the interval where the two functions coincide. Of course, the non-interiority Condition 2.1 is assured if w is strictly concave, but there is a broader set of functions satisfying it.[13] What we achieve is the substitution of the hypothesis of strict concavity and differentiability by concavity and the non-interiority condition. Observe that the second is a *local* condition.

---

[12] The notation $\partial^+ w(c)$ stands for the upper subdifferential of the function w at the point c; see Rockafellar (1970) for detailed definitions. We know that $\partial^+ w(\cdot) \neq \phi$ due to the concavity of w and that $z > 0$ because w is strictly increasing.

[13] Any concave piecewise linear function that has a kink at b_σ/σ or any concave function that is strictly concave in an interval containing b_σ/σ would be relevant examples; (Khan and Piazza, 2009) provide further diagrammatic illustrations of such functions. Also see the text pertaining to Footnote 4 above for the desirability of such generalized functions.

Next, we define the following subsets of $\mathbb{R}^n_+$:

$$S_c = \{c \in \mathbb{R}^n_+ / c_\sigma \in S_f \quad \text{and } c_i = 0 \quad \text{for all } i \neq \sigma\}$$
$$V = \{x \in \Delta / x_i \in S_f \quad \text{for all } i \leq \sigma \quad \text{and } x_i = 0 \text{ for all } i > \sigma\} \tag{5}$$

As discussed in Khan and Piazza (2009), the following results are obtained without the non-interiority Condition 2.1.

Proposition 2.3. *The von Neumann facet is*

$$\{(x, x') \in \Omega / \delta(x, x') = 0\} = \{(x, x') \in \Omega / \lambda(x, x') \in S_c\}.$$

Remark 2.1. Given $x \in V$, consider the σ-periodic program from x where the harvest consist of all the trees of the σ-th age class. This particular program has zero accumulated value loss, hence $\mu(x) = 0$ iff $x \in V$.

The following lemma, analogous to Mitra and Wan (1986, Lemma 6.4), is arguably a basic result not only for the forestry model, but also more generally for the theory of intertemporal allocation of resources.

Lemma 2.2. *Every good program* $\{x(t)\}$ *is such that* $dist(x(t), V) \to 0$

Next, we present the optimality criteria we shall be working with.

Definition 2.7. A program $\{x^*(t)\}$ is *optimal* if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have

$$\limsup_{T\to\infty} \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) \leq 0$$

Definition 2.8. A program $\{x^*(t)\}$ is *maximal* if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have

$$\liminf_{T\to\infty} \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) \leq 0$$

If the non-interiority Condition 2.1 does not hold, we cannot assure the existence of an optimal program from any $x_0 \in \Delta$, but only that of a maximal program. This follows from Proposition 2.2 that assures the existence of a minimizer of the accumulated value loss function and the following result.

Proposition 2.4. *If* $\{x(t)\}$ *is a program from* x_0 *that minimizes the accumulated value loss* ($\sum_t \delta(x(t), x(t+1)) = \mu(x_0)$), *then* $\{x(t)\}$ *is a maximal program from* x_0.

Corollary 2.1. *Every maximal program is good. Hence it converges to the set V.*

Lemma 2.3. *The non-interiority* Condition 2.1 *holds iff* $\{\hat{x}\} = V$.

From now on, and throughout the rest of this work, we will assume that the non-interiority Condition 2.1 holds. Let us spell out two preliminary results that we have with this added hypothesis. First, we present a stronger version of Lemma 2.2,

Lemma 2.4. *Any good program* $\{x(t)\}$ *satisfies* $\lim_{t\to\infty} x(t) = \hat{x}$.

Second, the existence of an optimal program is assured by the following equivalence,

Theorem 2.2. *Let* $\{x(t)\}$ *be a program from* x_0. *If the non-interiority* Condition 2.1 *holds, the following conditions are equivalent:*

i. $\sum_{t=0}^{\infty} \delta(x(t), x(t+1)) = \mu(x(0))$
ii. $\{x(t)\}$ *is optimal.*
iii. $\{x(t)\}$ *is maximal.*

Next, we observe that the two basic results in Khan and Piazza (2010) can be generalized under the non-interiority Condition 2.1 to yield the following versions whose straightforward proofs we leave to the reader. These are the only new results (so to speak) in this section.

Theorem 2.3. *Let* $\epsilon > 0$. *If the non-interiority* Condition 2.1 *holds, there exists* $\gamma > 0$ *such that for each optimal program* $\{x(t)\}$ *satisfying* $\|x(0) - \hat{x}\|_\infty < \gamma$ *the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \text{ for all } t \geq 0$$

Theorem 2.4. *Let* $\epsilon > 0$. *If the non-interiority* Condition 2.1 *holds, there exists a natural number* T_0 *such that for each optimal program* $\{x(t)\}$ *the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \text{ for all } t \geq T_0$$

3. Principal results

We now present the principal results of this work. As emphasized in the introduction, both results are classical in that they pertain to the following situation: if the initial and terminal configurations of a forest are given, then with enough time at his or her disposal, the planner knows that every optimal (or approximately optimal) program ought to stay arbitrarily near the golden-rule configuration, during an arbitrarily large fraction of the total time. However, the less of an error that the planner is allowed in steering the forest configuration away from the golden-rule configuration, the larger the time horizon required to observe this behavior. We point out that while the forest remains arbitrarily close to the golden rule configuration, the timber yields are arbitrarily close to the maximally sustainable one.

For the rest of the paper, we will assume that the non-interiority Condition 2.1 holds. We introduce notation for the aggregate value of finite optimal programs. Let $z_0, z_f \in \Delta$, $T \geq 1$, and

$$U(z_0, T) = \sup \left\{ \sum_{t=0}^{T-1} w(bc(t)) / \{x(t)\}_{t=0}^{T} \text{ is a program from } z_0 \right\}, \tag{6}$$

$$U(z_0, z_f, 0, T) = \sup \left\{ \sum_{t=0}^{T-1} w(bc(t)) / \{x(t)\}_{t=0}^{T} \text{ is a program from } z_0 \text{ with } x(T) = z_f \right\}. \tag{7}$$

Whenever there is no program $\{x(t)\}_{t=0}^{T}$ such that $x(0) = z_0$ and $x(T) = z_f$, we shall assume as a matter of mathematical conventions that $U(z_0, z_f, 0, T) = -\infty$.

Theorem A. *Given $M > 0$ and $\epsilon > 0$ there exists $L \in \mathbb{N}$ such that for all $T > L$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M$$

we have

$$Card\{i \in [0, \ldots, T-1] : \|x(t) - \hat{x}\| > \epsilon\} \leq L.$$

Theorem B. *Let $\epsilon > 0$. Then there exist $L \in \mathbb{N}$ and $M > 0$ such that for all $T > 2L + n + \sigma$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M$$

there are τ_1, τ_2 such that $\tau_1 \in [0, L]$, $\tau_2 \in [T - \tau, T]$ and

$$\|x(t) - \hat{x}\| \leq \epsilon \text{ for all } t = \tau_1, \ldots, \tau_2.$$

Moreover, if $\|x(0) - \hat{x}\| \leq \epsilon/n^2$ then $\tau_1 = 0$.

In terms of basic conception, the results are classical in the sense that the term is delineated in Khan and Zaslavski (2010a). Thus, rather than the Samuelsonian triple limit alluded to in his Nobel Lecture (Samuelson, 1971), an interesting "quarter limit" is involved, and four separate considerations are quantified: the length of the time-horizon, the proximity to the golden-rule forest configuration, the length of time that is spent within this proximity, and a "degree of slack" in the attainment of the objective. For an index of proximity quantified by ϵ, and index of slack quantified by M, Theorem A furnishes a bound L such that for all time-horizon levels T greater than L, any M-optimal forest configuration lies within the golden-rule configuration for $(T - L)$ number of periods. Since L is independent of the time-horizon, the optimal configuration lies close to the golden-rule configuration most of the time. If the planner is not allowed to choose the degree of slack M, the result can be strengthened to guarantee that the time-periods spent in proximity to the turnpike are consecutive. This is formalized in Theorem B, the analogue (in terms of approximately optimal programs) of the time-honored strengthening of Radner's result by Nikaido; see Khan and Zaslavski (2010a) for references and further discussion in terms of the RSS model.

Next, we observe that Theorem 2.4 above also follows as a straightforward consequence of Theorem B. To see this, simply note that any optimal infinite-horizon program $\{x(t)\}$, when truncated to T periods, is an optimal T-period program with its own initial and terminal configurations, $x(0)$ and $x(T)$; and that the Theorem furnishes us with L independent of both T *and* these forest configurations. This alternative proof is of some methodological significance in that it shows that a result on (uniform) asymptotic stability of the golden-rule forest configuration follows from the turnpike result classically conceived, and thereby establishes the primacy of McKenzie's so called *middle* turnpike over his *late* turnpike. We point out that the primacy of the former over McKenzie's *early* turnpike is already established by the second statement of Theorem B.

Theorems A and B also allow the deduction of the following two corollaries pertaining to the "bunching" of two approximately optimal programs with the same initial and terminal stocks. The first is phrased without any reference to the maximally sustainable forest configuration, and the second allows this "bunching" to be contiguous in an (arbitrarily large) initial time-interval if the initial forest configuration is the golden-rule configuration.

Corollary 3.1. *Given $\epsilon > 0$ and $M \geq 0$, there exists $L \in \mathbb{N}$ such that for each $T > L$, and any two programs $\{x_a\}_{t=0}^{T}$ and $\{x_b\}_{t=0}^{T}$ satisfying*

$$x_a(0) = x_b(0) = x_0, \quad x_a(T) = x_b(T) = x_T, \quad \sum_{t=0}^{T-1} w(bc(t)) \geq U(x_0, x_T, 0, T) - M,$$

where $c(t)$ stands alternatively for $c_a(t)$ and $c_b(t)$, following inequality holds:

$$Card\{i \in [0, \ldots, T-1] : \|x_a(t) - x_b(t)\|_\infty > \epsilon\} \leq L.$$

Corollary 3.2. *Given $\epsilon > 0$, there exists $L \in \mathbb{N}$ and $M > 0$ such that for each $T > 2L + n + \sigma$, and any two programs $\{x_a\}_{t=0}^{T}$ and $\{x_b\}_{t=0}^{T}$ satisfying*

$$x_a(0) = x_b(0) = \hat{x}, \quad x_a(T) = x_b(T) = x_T, \quad \sum_{t=0}^{T-1} w(bc(t)) \geq U(x_0, x_T, 0, T),$$

where $c(t)$ stands alternatively for $c_a(t)$ and $c_b(t)$, there are τ_1, τ_2 such that $\tau_1 \in [0, \tau]$, $\tau_2 \in [T - L, T]$ and

$$\|x_a(t) - x_b(t)\|_\infty \leq \epsilon \text{ for all } t = \tau_1, \ldots, \tau_2.$$

If in addition, $x_a(0) = x_b(0) = \hat{x}$, then $\tau_1 = 0$.

The corollaries are straightforward consequences of Theorems A and B. The proofs of the latter are developed in the sequel in two stages: Sections 4 and 5 below present the basic footholds on which the proofs rest while Section 7 spells out the technical details. We conclude this section with two observations.

First, whereas the proofs are a testimony to the fact that the standard results of classical turnpike theory do not directly apply, its methods and basic constructions can be refashioned in Sections 4 and 5 to develop the argumentation. However, a rather natural alternative proof-procedure suggests itself. This is to work with the Fenchel biconjugate of the felicity function, use the non-interiority condition supplemented possibly by some additional conditions to guarantee its strictly concavity and smoothness, apply the standard theory, and then "push it down" to the generalized context of Theorems A and B. It remains an open question whether this alternative approach can be executed.[14]

The second observation concerns the explicit dependence of the time-horizon L furnished in Theorems A and B on the error ϵ, and in Theorem A, on the degree of slack M. In particular, one would like to know the impact of initial and terminal configurations of the forest, and their proximity to the golden-rule configuration, on the time to reach the latter configuration. We note that our method of proof does not allow such a sensitivity analysis of the time needed to reach the $\epsilon-$ neighborhood of the golden rule stock with respect to the initial and terminal condition, and such quantitative estimates remain a topic for further study. However, see the Remark after the proof of Theorem A below.[15]

4. Preliminary substantive results

In this section and the next we begin developing the technical arguments needed to prove the principal results of this work. The five propositions presented here develop intuition into the basic dynamics underlying the MW model, and even though the proofs are notationally somewhat complex, the essential ideas are simple.[16] We shall be going into some detail as regards the comparison with Khan and Zaslavski (2010a), but for readers not particularly interested in the comparison, let us simply observe here that despite their superficial resemblance, Propositions 6.1–6.4 in Khan and Zaslavski (2010a) do not furnish the precise estimates that are offered here in Propositions 4.1–4.4, and strictly speaking there is no analogue to them in the RSS model.

Proposition 4.1 is a basic result of the subject that given any two forest configurations, there exists a program of n time periods that allows the planner to move from one configuration to another, n being the number of ages at which a particular tree can be tracked.[17]

Proposition 4.1. *For every $z_0, z_f \in \Delta$, there exists a program $\{x(t)\}_{t=0}^{n}$ such that $x(0) = z_0$ and $x(n) = z_f$*

[14] We are indebted to an anonymous referee for this paragraph. He asked whether it is "possible to get main results as limit cases of the well known theory? It looks plausible that the non-interiority condition is exactly the condition that gives such a possibility by smoothing and "strictly concavifying" the original utility function."

[15] We are indebted to the question of an anonymous referee for this paragraph, and for the answer detailed in the Remark below.

[16] The reader could check her understanding of the basics of the model by trying to figure out the proofs for herself before looking at the ones presented in the Appendix A.

[17] A version of this result was first presented in Mitra (2005).

The following corollary presents a refinement to programs that are not of n time periods.

Corollary 4.1. *Let $z_0, z_f \in \Delta$. (i) If $T \geq n$, there exists a program $\{x(t)\}_{t=0}^{T}$ such that $x(0)=z_0$ and $x(T)=z_f$. (ii) If $T<n$, there exists a program $\{x(t)\}_{t=0}^{T}$ such that $x(0)=z_0$ and $x(T)=z_f$ iff $z_{0,i} \geq z_{f,i+T}$ for all $i=1, \ldots, T-n$.*

There is no presumption that the program whose existence is asserted in the above claims is optimal in any sense. We turn to finite optimal programs in the next two results. Proposition 4.2 claims that the average benefit obtained from a given initial configuration can get arbitrarily close to that obtained from the maximally sustainable timber yield, as the time horizon becomes large enough. Proposition 4.3 makes a similar claim when the terminal forest configuration is also specified.

Proposition 4.2. *For each $z \in \Delta$, and each $T \in \mathbb{N}$,*

$$U(z, T) \geq Tw(b\hat{c}) - \sigma w(b\hat{c}).$$

Proposition 4.3. *Given $z_0, z_f \in \Delta$ and $T \geq n$, we have*

$$U(z_0, z_f, 0, T) \geq Tw(b\hat{c}) - (n+\sigma)w(b\hat{c}). \tag{8}$$

If $T<n$ and there is a program $\{x(t)\}_{t=0}^{T}$ satisfying that $x(0)=z_0$ and $x(T)=z_f$ then inequality (8) also holds.

Next, we turn to a simple inequality that follows from the fact that the value-loss of any production plan is non-negative.

Proposition 4.4. *For every T and every program $\{x(t)\}_{t=0}^{T}$ the following inequality is satisfied:*

$$\sum_{t=0}^{T-1}[w(bc(t)) - w(b\hat{c})] \leq n\left(\frac{b_\sigma}{\sigma}\right) z \tag{9}$$

Our final result asserts that any finite program that is optimal with a particular level of approximation, has its sub-programs also optimal with respect to the same level of approximation. It is analogous to Khan and Zaslavski (2010a, Proposition 6.6), we refer the reader to this article for the proof.

Proposition 4.5. *Let $T \in \mathbb{N}$, $M>0$ and $\{x(t)\}_{t=0}^{T}$ be a program such that*

$$\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M.$$

Then for all S_1 and S_2, $0 \leq S_1 < S_2 < T$, we have

$$\sum_{t=S_1}^{S_2-1} w(bc(t)) \geq U(x(S_1), x(S_2), S_1, S_2) - M$$

5. Four substantive lemmas

With these preliminary results out of the way, we can turn to the deeper substance of the argumentation. As in Khan and Zaslavski (2010a), it revolves around the four footholds presented below as Lemmas 5.1–5.4: the *visiting* lemma, the *stability* lemma, the *value-loss* lemma and the *aggregate value-loss* lemma. However, before we take each in turn, it is worth elaborating on what was already emphasized in the introduction: that even though these results are inspired by Lemmas 7.1 to 7.4 in Khan and Zaslavski (2010a) for the RSS model, the particularities of the MW model bring in analytical difficulties of their own that need to be surmounted. Briefly put, in the RSS model a unit amount of labor is to be allocated to the production of a single consumption good and to the production of n types of machines; whereas in the MW model, a unit amount of land is to be parcelled out between the cultivation of trees of n possible ages. Thus in one, the stock variable is an element of $\mathbb{R}^n$ with a well-defined order on it; whereas in the other, it is a probability measure with a finite support that can be represented as a point in the simplex in $\mathbb{R}^n$, with no clear order, and therefore little possibility of formalizing notions either of *i* naction or of *free disposal.* [18] All this renders the proof of a result in one inapplicable to that of the other, and adds to their considerable complication. This is especially true of the Lemmas 5.3 and 5.4 below, the Radner (1961) *value-loss* and *aggregate value-loss* lemmas. Note also the considerable sharpening of the conclusion of Lemma 5.2 and Corollary 5.1 relative to their RSS counterparts in Khan and Zaslavski (2010a).

[18] This point of view as regards the Mitra–Wan tree-farm is original to Khan and Piazza (2010).

Lemma 5.1. *Given $M > 0$ and $\epsilon > 0$, there exists $\tau \in \mathbb{N}$ such that for each program $\{x(t)\}_{t=0}^{\tau}$ satisfying*

$$\sum_{t=0}^{\tau-1} w(bc(t)) \geq \tau w(b\hat{c}) - M,$$

there exists $t \in [0, \tau]$ such that $\|x(t) - \hat{x}\| \leq \epsilon$.

Lemma 5.2. *Let the program $\{x(t)\}_{t=0}^{n}$ be such that*

$$\delta(x(t), x(t+1)) = 0 \quad \text{for } t = 0, \ldots, n-1. \tag{10}$$

then $x(t) = \hat{x}$ for all $t \in [0, \sigma]$.

Corollary 5.1. *Given $\epsilon > 0$, there exist $\gamma > 0$ such that for each program $\{x(t)\}_{t=0}^{n}$ satisfying $\delta(x(t), x(t+1)) < \gamma$ for $t = 0, \ldots, n-1$, we have $\|x(t) - \hat{x}\| < \epsilon$ for all $t = 0, \ldots, \sigma$.*

Lemma 5.3. *Given $\epsilon > 0$, there exists $\gamma > 0$ such that for each $T \in \mathbb{N}$, and each program $\{x(t)\}_{t=0}^{T}$ satisfying $\|x(0) - \hat{x}\| < \epsilon/n$, $\|x(T) - \hat{x}\| < \epsilon/n$ and $\delta(x(t), x(t+1)) < \gamma$ for all $t = 0, \ldots, T-1$, we have*

$$\|x(t) - \hat{x}\| < \epsilon \text{ for all } t = 0, \ldots, T.$$

Lemma 5.4. *Given $\epsilon > 0$, there exist $\gamma > 0$ and $M > 0$ such that for each $T \geq n + \sigma$, and each program $\{x(t)\}_{t=0}^{T}$ satisfying $\|x(0) - \hat{x}\| \leq \gamma$, $\|x(T) - \hat{x}\| \leq \gamma$ and $\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M$, we have*

$$\sum_{t=0}^{T-1} \delta(x(t), x(t+1)) < \epsilon.$$

For a detailed discussion and interpretation of these results, the reader is referred to Khan and Zaslavski (2010a, Section 4).

6. Concluding remarks

We now conclude the non-technical part of this work by delineating three directions in which the results demand extension and further investigation.

The first of these is the rather immediate question as to how much of the theory can be salvaged when the non-interiority Condition 2.1 does *not* hold? Since this condition is necessary *and* sufficient for asymptotic convergence of good programs, and it is easy to provide examples of periodic optimal programs when it does not hold, perhaps the obvious answer is simply: none of it. However, the question clearly deserves another less-facile look. The fact that infinite-horizon optimal optimal programs converge to the von-Neumann facet is a basic result of the subject, and surely what is true of asymptotic convergence could also possibly be true of classical turnpike theory.

This work, along with that of Khan and Zaslavski (2010a), has taken classical turnpike theory (circumscribed as it is by assumptions of uniform strict concavity, and on occasion, differentiability, of felicity functions) and extended it to concave functions that are not necessarily differentiable. In terms of a second direction, one is then naturally led to ask whether one can relax the concavity (and continuity) assumption itself?

This question has not been posed so far in the capital-theory literature, but if mathematical economics is to justify itself as an intellectually worthwhile activity, surely that justification must revolve in part on the acceptance of each new result leading to the pursuit of questions that would not have been considered even remotely feasible before it.

These two directions stems directly from the two theorems reported here; the third is rather more overarching. A subtext of this entire work is the (somewhat uneasy) relationship between the RSS and MW models, with the relative difficulties of one being matched by those of another, and presenting inevitable analytical trade-offs. This tension clearly asks for a move towards a synthesis that obtains both models as special cases, and provides a non-trivial extension to what is now associated with the names of Gale and McKenzie and is frequently referred to as the general theory of intertemporal allocations of resources.

Appendix A. Technicalities of proofs

Proof of Proposition 4.1. We propose the following program $\{x(t)\}_{t=0}^{n}$,

$$
\begin{array}{rclrcl}
x(0) & = & z_0 & c(0) & = & z_0 \\
x(1) & = & (1, 0, \ldots, 0) & c(1) & = & (\sum_{i=1}^{n-1} z_{f,i},\, 0, \ldots, 0) \\
x(2) & = & (\sum_{i=1}^{n-1} z_{f,i},\, z_{f,n},\, 0, \ldots, 0) & c(2) & = & (\sum_{i=1}^{n-2} z_{f,i},\, 0, \ldots, 0) \\
 & \vdots & & \vdots & \vdots & \vdots \\
x(j) & = & (\sum_{i=1}^{n-j+1} z_{f,i},\, z_{f,n-j+2}, \ldots, z_{f,n-1},\, z_{f,n},\, \underbrace{0, \ldots, 0}_{n-j}) & c(j) & = & (\sum_{i=1}^{n-j} z_{f,i},\, 0, \ldots, 0) \\
\vdots & \vdots & \vdots & & & \\
x(n) & = & z_f \quad \square & & &
\end{array}
$$

Proof of Corollary 4.1. (i) First consider *any* program $\{x(t)\}_{t=0}^{T-n}$ from z_0. The proposition above tells us that there is a program $\{\bar{x}(t)\}_{t=0}^{n}$ where $\bar{x}(0) = x(T-n)$ and $\bar{x}(n) = z_f$. The proof follows by defining $x(t+T-n) = \bar{x}(t)$ for $t = 1, \ldots, n$. (ii) It follows easily from the definition of program. □

Proof of Proposition 4.2. If $T \leq \sigma$ then the proposition follows directly because $U(z, T) \geq 0 \geq Tw(b\hat{c}) - \sigma w(b\hat{c})$.
If $T > n$, consider the following program $\{x(t)\}_{t=0}^{\infty}$ from z

$$
\begin{array}{rclrcl}
x(0) & = & z & c(0) & = & z \\
x(1) & = & (1, 0, \ldots, 0) & c(1) & = & \left(\frac{\sigma-1}{\sigma}, 0, \ldots, 0\right) \\
x(2) & = & \left(\frac{\sigma-1}{\sigma}, \frac{1}{\sigma}, 0, \ldots, 0\right) & c(2) & = & \left(\frac{\sigma-2}{\sigma}, 0, \ldots, 0\right) \\
\vdots & \vdots & & \vdots & \vdots & \\
x(j) & = & \left(\frac{\sigma-j+1}{\sigma}, \underbrace{\frac{1}{\sigma}, \ldots, \frac{1}{\sigma}}_{j-1}, \underbrace{0, \ldots, 0}_{n-j}\right) & c(j) & = & \left(\frac{\sigma-j}{\sigma}, 0, \ldots, 0\right) \\
\vdots & \vdots & & \vdots & \vdots & \\
x(t) & = & \hat{x} \quad \text{for all } t \geq \sigma & c(t) & = & \hat{c} \quad \text{for all } \sigma \leq t \leq T
\end{array}
$$

And we deduce

$$U(z, T) \geq \sum_{t=0}^{T-1} w(bc(t)) \geq \sum_{t=\sigma}^{T-1} w(bc(t)) = (T-\sigma)w(b\hat{c}) \quad \square$$

Proof of Proposition 4.3. If $T \geq n$, by Proposition 4.2 we know that there is a program $\{x(t)\}_{t=0}^{T-n}$ from z_0 such that

$$\sum_{t=0}^{T-n-1} w(bc(t)) \geq (T-n)w(b\hat{c}) - \sigma w(b\hat{c})$$

and by Proposition 4.1, there is a program $\{x(t)\}_{t=T-n}^{T}$ from $x(T-n)$ such that $x(T)=z_f$. Concatenating the two, we obtain the program $\{x(t)\}_{t=0}^{T}$ from z_0 such that $x(T)=z_f$ and we can deduce

$$U(z_0, z_f, 0, T) \geq \sum_{t=0}^{T-1} w(bc(t)) \geq \sum_{t=0}^{T-n-1} w(bc(t)) \geq (T-n)w(b\hat{c}) - \sigma w(b\hat{c}) = Tw(b\hat{c}) - (n+\sigma)w(b\hat{c}).$$

If $T<n$, then the existence of the program $\{x(t)\}_{t=0}^{T}$ assures that $U(z_0, z_f, 0, T) \geq 0$ and observing that $(T-n-\sigma)w(bc(t)) < 0$ the proof follows. □

Proof of Proposition 4.4. We know that $\delta(x(t), x(t+1)) \geq 0$, hence

$$0 \leq \sum_{t=0}^{T-1} \delta(x(t), x(t+1)) = \sum_{t=0}^{T-1} \left[w\left(\frac{b_\sigma}{\sigma}\right) - w(bc(t)) \right] + \hat{p}(x(0) - x(T))$$

Using the definition of Δ we deduce that

$$\sum_{t=0}^{T-1} w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \leq \|\hat{p}\| \, \|x(0) - x(T)\| \leq \|\hat{p}\| = n\frac{b_\sigma}{\sigma} z. \quad \square \tag{11}$$

Proof of Lemma 5.1. Let us assume the contrary: for each $k \in \mathbb{N}$ there exists a program $\{x^k(t)\}_{t=0}^{k}$ such that

$$\|x^k(t) - \hat{x}\| > \epsilon \text{ and } \sum_{t=0}^{k-1} w(bc^k(t)) \geq kw\left(\frac{b_\sigma}{\sigma}\right) - M \tag{12}$$

Let $M' = n\frac{b_\sigma}{\sigma} z > 0$, by (9) we know that every program fulfills $\sum_{t=0}^{T-1} w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \leq M'$.

Given any $s<k$, by combining the two previous inequalities we deduce

$$\sum_{t=0}^{s-1} \left[w(bc^k(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \right] = \sum_{t=0}^{k-1} \left[w(bc^k(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \right] - \sum_{t=s}^{k-1} \left[w(bc^k(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \right] \geq -(M + M') \tag{13}$$

By extracting a subsequence and a diagonalization process we obtain that there exist a strictly increasing sequence of natural numbers $\{k_j\}_{j=1}^{\infty}$ and a sequence $\{x^*(t)\}_{t \in \mathbb{N}}$ such that

$$x^{k_j}(t) \to x^*(t) \text{ when } j \to \infty \text{ for all } t \geq 0$$

It is not difficult to see that $\{x^*(t)\}_{t \in \mathbb{N}}$ is a program. From (13) we deduce that for every natural number s, $\sum_{t=0}^{s-1} w(bc^*(t)) - sw(b_\sigma/\sigma) \geq -M - M'$, meaning that $\{x^*(t)\}_{t \in \mathbb{N}}$ is a good program. Then Lemma 2.2 implies that

$$\|x^*(t) - \hat{x}\| \to 0 \text{ when } t \to \infty.$$

On the other hand, it follows from (12) and the definition of $x^*(t)$ that

$$\|x^*(t) - \hat{x}\| > \epsilon \text{ for all } t$$

and the contradiction proves the lemma. □

Proof of Lemma 5.2. By Lemma 2.3, we know that $c(t) \in S_c$ for $0 \leq t < n$ which implies that $x_i(\sigma)=0$ for all $i>\sigma$. Indeed, if there was $j>\sigma$ such that $x_j(\sigma)>0$, then we would have $x_n(n+\sigma-j)>0$ and $c(n+\sigma-j) \notin S_c$.

From the above we know that

$$x(\sigma) = (c_\sigma(\sigma-1), c_\sigma(\sigma-2), \ldots, c_\sigma(0), 0 \ldots, 0)$$

where $c_\sigma(i) \in S_f$. The area balance together with Condition 2.1 implies $x(\sigma) = \hat{x}$.

Finally, it is easy to see that $x(t+1) = \hat{x}$ and $\delta(x(t), x(t+1)) = 0$ imply $x(t) = \hat{x}$ and then the proposition follows by backwards induction. □

Proof of Corollary 5.1. Suppose, contrary to our claim, that for every k there is $\{x^k(t)\}_{t=0}^{n}$ such that $\delta(x^k(t), x^k(t+1)) \leq 1/k$ and there is $t_k \in [0, \ldots, \sigma]$ satisfying $\|x^k(t_k) - \hat{x}\| \geq \epsilon$. As $\{t_k\} \subseteq [0, \ldots, \sigma]$ there must be at least one value $t_0 \in [0, \ldots, \sigma]$ such that $t_k = t_0$ infinitely many times. We extract a subsequence $\{k_j\}$ such that $t_{k_j} = t_0$ and hence $\|x^{k_j}(t_0) - \hat{x}\| \geq \epsilon$. By extracting a subsequence from $\{k_j\}$ (to simplify the notation, we denote this subsequence also by $\{k_j\}$), we obtain that there is a sequence $\{x^*(t)\}$ such that

$$x^{k_j}(t) \to x^*(t) \quad \text{for all } t = 0, \ldots, n.$$

It is easy to see that $\{x^*(t)\}$ is a program and by the continuity of $\delta(\cdot,\cdot)$ we have that $\delta(x^*(t), x^*(t+1))=0$ for all $t=0,\ldots,n$ and the lemma above implies that $x^*(t)=\hat{x}$ for all $t=0,\ldots,\sigma$.

On the other hand, $\|x^{k_j}(t_0)-\hat{x}\|\geq\epsilon$ for all k_j hence $\|x^*(t_0)-\hat{x}\|\geq\epsilon$ and a contradiction arises proving the corollary. □

Proof of Lemma 5.3. We divide the proof into two parts: $T<n$ and $T\geq n$.

Case $T<n$: A procedure similar to the one on the corollary above allows to affirm that given ϵ there is γ_1 such that: $\delta(x, x')<\gamma_1$ implies $\operatorname{dist}(\lambda(x,x'), S_c)<\epsilon_1=\epsilon/(n2^n)$.[19]

Although the computation is quite cumbersome, the argument of the proof is based in a simple idea: if the distances $\|x(0)-\hat{x}\|$ and $\|x(T)-\hat{x}\|$ are small and the harvesting policy is similar to the periodic program harvesting, then the state of the forest cannot go far from $\hat{x}$ (in less than n steps) without making a large value loss in at least one step.

To see that $\|x(t)-\hat{x}\|<\epsilon$ for all $t=1,\ldots,T-1$ we start bounding the value of $x_i(t)$ for all $i=\sigma+1,\ldots,n$ and after that we bound $||x_i(t)-(1/\sigma)||$ for all $i=1,\ldots,\sigma$.

First consider $i>\sigma$. In this case, we can express $x_i(t)$ as a linear combination of $x_{i+t-T}(T)$ or $x_n(t+n-i)$ and the harvests between t and T or $t+n-i$ (that are controlled by ϵ_1) to deduce that

$$x_i(t)<\frac{\epsilon}{2^n}+n\epsilon_1 \text{ for all } i>\sigma. \tag{14}$$

We need to divide the study into two cases:

1. Case $i-t+T\leq n$,

$$x_i(t)=x_{i+T-t}(T)+\sum_{j=0}^{T-t-1}c_{i+j}(t+j)<\frac{\epsilon}{2^n}+(T-t)\epsilon_1<\frac{\epsilon}{2^n}+n\epsilon_1<\epsilon$$

2. Case $i-t+T>n$,

$$x_i(t)=x_n(t+n-i)+\sum_{j=0}^{n-i-1}c_{i+j}(t+j)=\sum_{j=0}^{n-i}c_{i+j}(t+j)<n\epsilon_1<\epsilon$$

To deal with the i th age class when $i\leq\sigma$ we start by proving that

$$\text{if } \|x(t)-\hat{x}\|<\epsilon_2 \text{ then } |x_i(t+1)-\frac{1}{\sigma}|<n\epsilon_1+2\epsilon_2 \quad \text{for all } i=1,\ldots,\sigma. \tag{15}$$

Case $2\leq i\leq\sigma$,

$$|x_i(t+1)-\frac{1}{\sigma}|=|x_{i-1}(t)-c_{i-1}(t)-\frac{1}{\sigma}|\leq|c_{i-1}(t)|+|x_{i-1}(t)-\frac{1}{\sigma}|<\epsilon_1+\epsilon_2$$

Case $i=1$,

$$\begin{aligned} x_1(t+1)=\sum_{i=1}^{n}c_i(t) &\Rightarrow c_\sigma(t)\leq x_1(t+1)\leq(n-1)\epsilon_1+c_\sigma(t)\\ \Rightarrow |x_1(t+1)-\frac{1}{\sigma}| &\leq (n-1)\epsilon_1+|x_\sigma(t)-x_{\sigma+1}(t+1)-\frac{1}{\sigma}|\\ &\leq (n-1)\epsilon_1+|x_\sigma(t)-\frac{1}{\sigma}|+|x_{\sigma+1}(t+1)|<(n-1)\epsilon_1+2\epsilon_2 \end{aligned}$$

Repeated application of (14) and (15) yields,

$$\begin{aligned} \|x(0)-\hat{x}\|<\frac{\epsilon}{2^n} &\Rightarrow \|x(1)-\hat{x}\|<2\frac{\epsilon}{2^n}+n\epsilon_1\\ &\Rightarrow \|x(2)-\hat{x}\|<2(2\frac{\epsilon}{2^n}+n\epsilon_1)+n\epsilon_1\\ &\quad\vdots\\ &\Rightarrow \|x(T-1)-\hat{x}\|<2^{T-1}\frac{\epsilon}{2^n}+(2^{T-1}-1)n\epsilon_1 \end{aligned}$$

and thus $\|x(t)-\hat{x}\|<\epsilon$ for all $t=1,\ldots,T-1$.

[19] $S_c=\{c\in\mathbb{R}^n_+ : c_\sigma\in S_f \text{ and } c_i=0 \text{ for all } i\neq\sigma\}$

Case $T \geq n$: Corollary 5.1 states that there is γ_2 such that for every program $\{x(t)\}_{t=0}^{n}$ satisfying $\delta(x(t), x(t+1)) < \gamma_2$ for all $t < n$, we have $\|x(t) - \hat{x}\| < \epsilon/2^n$ for all $t = 0, \ldots, \sigma$. Apply this result to the programs $\{x(t+i)\}_{t=0}^{n}$ with $i = 0, \ldots, T - \sigma$ to conclude that $\|x(t) - \hat{x}\| < \epsilon/2^n < \epsilon$ for $t = 0, \ldots, T - n + \sigma$. Afterwards, apply part 1, to conclude that during the last $n - \sigma$ steps the state also fulfills: $\|x(t) - \hat{x}\| < \epsilon$ for $t = T - n + \sigma, \ldots, n$.

Take $\gamma = \min\{\gamma_1, \gamma_2\}$ and the lemma follows. □

Proof of Lemma 5.4. First observe that the hypothesis $\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M$ implies

$$M \geq \sum_{t=0}^{T-1} w(bc'(t)) - w(bc(t))$$

for any program $\{x'(t)\}$ such that $x'(0) = x(0)$ and $x'(T) = x(T)$. We can then find a bound of the accumulated value loss of $\{x(t)\}$ related to the accumulated value loss of $\{x'(t)\}$,

$$\begin{aligned} \sum_{t=0}^{T-1} \delta(x(t), x(t+1)) &= \sum_{t=0}^{T-1} \left[w\left(\frac{b_\sigma}{\sigma}\right) - w(bc(t)) \right] - \hat{p}(x(T) - x(0)) \\ &\leq M + \sum_{t=0}^{T-1} \left[w\left(\frac{b_\sigma}{\sigma}\right) - w(bc'(t)) \right] - \hat{p}(x'(T) - x'(0)) \\ &= M + \sum_{t=0}^{T-1} \delta(x'(t), x'(t+1)) \end{aligned}$$

Now, taking $M \leq \epsilon/2$ it suffices to prove that there is a program $\{x'(t)\}$ as above, yielding an accumulated value loss smaller than $\epsilon/2$. We build such a program to prove its existence. Given $T \geq n + \sigma$, we look for a program such that:

$$\begin{cases} \delta(x'(t), x'(t+1)) < \dfrac{\epsilon}{2(n+\sigma)} & t = 0, \ldots, \sigma - 1 \\ x'(t) = \hat{x} & t = \sigma, \ldots, T - n \\ \delta(x'(t), x'(t+1)) < \dfrac{\epsilon}{2(n+\sigma)} & t = T - n, \ldots, T \end{cases}$$

Of course, we have $\delta(x'(t), x'(t+1)) = 0$ for all $t = \sigma, \ldots, T - n - 1$.

Let γ_1 be such that $\delta(x'(t), x'(t+1)) < \frac{\epsilon}{2(n+\sigma)}$ if $\|(x'(t), x'(t+1)) - (\hat{x}, \hat{x})\| \leq \gamma_1$. Take $\gamma = \min\left\{\frac{\gamma_1}{n}, \frac{1}{n\sigma}\right\}$.

We start by dealing with the first σ elements of the program. We refer the reader to the proof of Khan and Piazza (2010, Proposition 5.5) where it is seen that for any $x(0)$ satisfying $\|x(0) - \hat{x}\| \leq \gamma$ there is a program $\{x'(t)\}_{t=0}^{\sigma}$ from $x(0)$ such that

$$\begin{cases} \|x'(t) - \hat{x}\| < \gamma_1, & t = 1, \ldots, \sigma - 1 \\ x'(\sigma) = \hat{x} \end{cases} \tag{16}$$

implying that $\delta(x'(t), x'(t+1)) \leq \frac{\epsilon}{2(n+\sigma)}$ for all $t < \sigma$.

To finish the proof, we need to find the last $n+1$ elements of the program $\{x'(t)\}$ satisfying the following conditions

$$\begin{cases} x'(T-n) = \hat{x} \\ \|x'(t) - \hat{x}\| < \gamma_1, & t = T - n + 1, \ldots, T - 1 \\ x'(T) = x(T) \end{cases} \tag{17}$$

We refer the reader to the proof of Khan and Piazza (2010, Lemma 6.2) where the program $\{x'(t)\}_{t=T-n}^{T}$ is built. □

We now turn to the proof of Theorem A. The idea of the proof is to use Lemma 5.3 to bound the difference $\|x(t) - \hat{x}\|$. In general, it will not be possible to apply this lemma to the whole interval $[0, T]$. To overcome this difficulty we divide $[0, T]$ in conveniently chosen subintervals of bounded lengths, so that the lemma will be valid in all but a finite number of subintervals, where this finite number depends only on M and ϵ.

Proof of Theorem A. Given ϵ, by Lemma 5.3 there is γ such that for each $\{x(t)\}_{t=0}^{T}$ satisfying

$$\|x(0) - \hat{x}\| < \epsilon/2^n, \ \|x(T) - \hat{x}\| < \epsilon/2^n \text{ and } \delta(x(t), x(t+1)) < \gamma \text{ for all } t \in [0, \ldots, T-1] \tag{18}$$

we have

$$\|x(t) - \hat{x}\| < \epsilon \text{ for all } t \in [0, \ldots, T] \tag{19}$$

Given a program $\{x(t)\}$ satisfying the hypothesis and taking S, τ such that $0 \le S \le S+\tau \le T$ we can use Proposition 4.5 to obtain

$$\sum_{t=S}^{S+\tau-1} w(bc(t)) \ge U(x(S), x(S+\tau), S, S+\tau) - M \tag{20}$$

and Proposition 4.3 to deduce

$$U(x(S), x(S+\tau), S, S+\tau) - M \ge \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M \tag{21}$$

From the above and Lemma 5.1 it follows that there is $\bar{\tau}$ such that

$$\text{for any } S \in [0, T-\bar{\tau}] \text{ there is } t \in [S, S+\bar{\tau}] \text{ such that } \|x(t)-\hat{x}\| < \epsilon/2^n \tag{22}$$

We next divide the interval $[0, T]$ in subintervals $[t_i, t_{i+1}]$ with $i = 0, \ldots, K$ where $t_0 = 0$, $t_K = T$ and

$$\bar{\tau} \le (t_i - t_{i-1}) \le 2\bar{\tau} \text{ and } \|x(t_i)-\hat{x}\| < \epsilon/2^n \text{ for all } i = 1, \ldots, K-1,$$

using the following algorithm: by (22) there is $t_1 \in [\bar{\tau}, 2\bar{\tau}]$ such that $\|x(t_1)-\hat{x}\| < \epsilon/2^n$. Using (22) again we know that there exists $t_2 \in [t_1+\bar{\tau}, \ldots, t_1+2\bar{\tau}]$ such that $\|x(t_2)-\hat{x}\| < \epsilon/2^n$. We proceed inductively defining

$$t_{i+1} \in [t_i+\bar{\tau}, t_i+2\bar{\tau}] \text{ with } \|x(t_{i+1})-\hat{x}\| < \epsilon/2^n$$

We repeat this step until we obtain $(t_{K-1}+2\bar{\tau}) \ge T$, then we set $t_K = T$ and the construction of the sequence is finished.

For every $i = 1, \ldots, K-2$, we can apply Lemma 5.3 whenever

$$\sum_{t=t_i}^{t_{i+1}-1} \delta(x(t), x(t+1)) < \gamma \tag{23}$$

to affirm that $\|x(t)-\hat{x}\| < \epsilon$ for all $t \in [t_i, t_{i+1}-1]$. We claim that there are $k \le 2+\gamma^{-1}[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z + M]$ subintervals *not* fulfilling (23). Indeed, denote by $\mathcal{K} \subseteq [1, \ldots, K-2]$ the set of indexes such that $\sum_{t=t_i}^{t_{i+1}-1} \delta(x(t), x(t+1)) \ge \gamma$, it is easily seen that

$$\begin{aligned}\sum_{t=0}^{T-1} \delta(x(t), x(t+1)) &= \sum_{k=0}^{K-1} \sum_{t=t_i}^{t_{i+1}-1} \delta(x(t), x(t+1)) \\ &\ge \sum_{k \in \mathcal{K}} \sum_{t=t_i}^{t_{i+1}-1} \delta(x(t), x(t+1)) \ge \gamma \,\mathrm{Card}\,\{\mathcal{K}\}.\end{aligned}$$

On the other hand we know that

$$\begin{aligned}\sum_{t=0}^{T-1} \delta(x(t), x(t+1)) &= \sum_{t=0}^{T-1} [w(b\hat{c}) - w(bc(t))] + \hat{p}(x(0) - x(T)) \\ &\le Tw(b\hat{c}) - U(x(0), x(T), 0, T) + \hat{p}(x(0) - x(T)) + M \\ &\le (n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z + M\end{aligned}$$

Combining the last two inequalities we get $\mathrm{Card}\,\{\mathcal{K}\} \le \gamma^{-1}[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z]$ and it follows that

$$\mathrm{Card}\,\{t = [0, \ldots, T] \text{ such that } \|x(t)-\hat{x}\| > \epsilon\} \le 2\bar{\tau}\left\{2+\gamma^{-1}[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z + M]\right\}.$$

Set $L = 2\bar{\tau}\{2+\gamma^{-1}[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z + M]\}$ and the theorem follows. □

Remark. A careful examination of the proof of Theorem A above allows us to see that the initial and terminal configurations play a rather unimportant role on the bound found for this time of convergence:

$$L = 2\bar{\tau}\{\mathbf{2} + \gamma^{-1}[(n+\sigma)w(b\hat{c}) + n(b_\sigma/\sigma)z + M]\}.$$

However, if we could assure a priori that both the initial and terminal configuration are within ϵ distance from the golden rule stock, our new bound would be smaller as the term "**2**" would be eliminated and L would reduce to

$$\tilde{L} = 2\bar{\tau}\{\gamma^{-1}[(n+\sigma)w(b\hat{c}) + n(b_\sigma/\sigma)z + M]\}. \quad \square$$

Theorem B aims for a stronger condition than Theorem A: not only $\|x(t)-\hat{x}\| < \epsilon$ must hold for most of the time stages, but these time stages must be consecutive, i.e., violations to the condition $\|x(t)-\hat{x}\| < \epsilon$ (if any) can only occur during the initial time stages or the last ones. We use again Lemma 5.3 to bound the difference $\|x(t)-\hat{x}\|$, but to apply it to an interval almost as large as $[0, T]$ we have to pay the price that we cannot choose the parameter M but that its value will be determined in keeping with the needs of the proof.

Proof of Theorem B. By Lemma 5.3 we know that given $\epsilon > 0$ there is $\epsilon_1 > 0$ such that for any program $\{x(t)\}_{t=\tau_1}^{\tau_2}$ satisfying

$$\|x(\tau_1)-\hat{x}\| \le \epsilon/2^n,\ \|x(\tau_2)-\hat{x}\| \le \epsilon/2^n,\ \text{and}\ \delta(x(t), x(t+1)) \le \epsilon_1 \text{ for all } t = \tau_1, \ldots, \tau_2 - 1 \tag{24}$$

the following inequality holds

$$\|x(t)-\hat{x}\| \le \epsilon \text{ for all } t = \tau_1, \ldots, \tau_2 \tag{25}$$

In order to bound $\delta(x(t), x(t+1))$ we use Lemma 5.4 which states that given $\epsilon_1 > 0$ there are $\gamma > 0$ and $M > 0$ such that for every τ_1 and τ_2 (satisfying $\tau_2 - \tau_1 \ge n + \sigma$) and every program fulfilling

$$\|x(\tau_1)-\hat{x}\| \le \gamma,\ \|x(\tau_2)-\hat{x}\| \le \gamma,\ \text{and} \sum_{t=\tau_1}^{\tau_2-1} w(bc(t)) \ge U(x(\tau_1), x(\tau_2), \tau_1, \tau_2) - M \tag{26}$$

we have that $\sum_{t=\tau_1}^{\tau_2-1} \delta(x(t), x(t+1)) \le \epsilon_1$, implying directly that $\delta(x(t), x(t+1)) \le \epsilon_1$ for all $t = \tau_1, \ldots, \tau_2 - 1$.

We proceed now to prove the existence of

$$\tau_1 \in [0, \tau] \text{ and } \tau_2 \in [T-\tau, T] \text{ such that } \|x(\tau_i)-\hat{x}\| < \gamma. \tag{27}$$

Let $M_1 = M + (n+\sigma)w(b\hat{c})$, Lemma 5.1 states that there is τ such that if the program $\{x(t)\}_{t=0}^{\tau}$ satisfies

$$\sum_{t=0}^{\tau-1} w(bc(t)) \ge \tau w(b\hat{c}) - M_1 \text{ then there is } t \in [0, \tau] \text{ such that } \|x(t)-\hat{x}\| < \gamma \tag{28}$$

By propositions 4.3 and 4.5 we know that

$$\sum_{t=0}^{\tau-1} w(bc(t)) \ge U(x(0), x(\tau), 0, \tau) - M \ge \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M$$

$$\sum_{t=T-\tau}^{T-1} w(bc(t)) \ge U(x(t-\tau), x(T), T-\tau, T) - M \ge \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M$$

and the existence of $\tau_1 \in [0, \tau]$ and $\tau_2 \in [T-\tau, T]$ such that $\|x(\tau_i)-\hat{x}\| < \gamma$ is assured.

Now Proposition 4.5 tells us that for any program fulfilling the hypothesis, and for any $0 \le \tau_1 < \tau_2 \le T$, $\sum_{\tau_1}^{\tau_2-1} w(bc(t)) \ge U(x(\tau_1), x(\tau_2), \tau_1, \tau_2) - M$. This fact together with (27) proves that (26) holds if $T \ge 2\tau + n + \sigma$ and then $\delta(x(t), x(t+1)) \le \epsilon_1$ for all $t = \tau_1, \ldots, \tau_2 - 1$. Finally, this proves that (24) holds and we conclude as desired. □

References

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

Gale, D., 1970. Nonlinear duality and qualitative properties of optimal growth. In: Abadie, J. (Ed.), Integer and Nonlinear Programming. North Holland, Amsterdam, pp. 309–319.

Kant, S., Berry, R.A. (Eds.), 2005. Economics, Sustainability, and Natural Resources: Economics of Sustainable Forest Management, vol. 1 and 2. Springer, Berlin.

Khan, M.A., Chapter 2, 39–65 in Kant and Berry (2005) 2005. Intertemporal Ethics, Modern Capital Theory and the Economics of Forestry.

Khan, M.A., Piazza, A., 2010. On uniform convergence of undiscounted optimal programs in the Mitra–Wan forestry model: the strictly concave case. International Journal of Economic Theory 6, 57–76.

Khan M.A., Piazza A., 2009. On the Mitra–Wan Forestry Model: A Unified Analysis. *Publicación Técnica CMM* No. 226. Paper presented at *Workshop in Economic Dynamics*, National University of Singapore, July 30 to August 1, 2009.

Khan, M.A., Zaslavski, A.J., 2010a. On two classical turnpike results for the Robinson-Solow-Srinivasan (RSS) model. Advances in Mathematical Economics 13, 47–97.

Khan, M.A., Zaslavski, A.J., 2010b. On locally optimal programs in the Robinson-Solow-Srinivasan model. Journal of Economics 99, 65–92.

Majumdar, M., Mitra, T., 1994. Periodic and chaotic programs of optimal intertemporal allocation in an aggregative model with wealth effects. Economic Theory 4, 677–688.

McKenzie, L., 1976. Turnpike theory. Econometrica 44, 841–865.

Mitra, T., Chapter 5, 137–173 in Kant and Berry (2005) 2005. Intergenerational Equity and the Forest Management Problem.

Mitra, T., 2006. Notes on a model of forestry: the undiscounted case. mimeo, Cornell University.

Mitra, T., Wan, H.W., 1985. Some theoretical results on the economics of forestry. Review of Economics Studies LII, 263–282.

Mitra, T., Wan, H.W., 1986. On the Faustmann solution to the forest management problem. Journal of Economic Theory 40, 229–249.

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

Rockafellar, R.T., 1970. Convex Analysis. Princeton University Press, Princeton. Reprint: 1997.

Rosser Jr., J.B., Chapter 9, 191–206 in Kant and Berry (2005) 2005. Complexities of Dynamic Forest Management Policies.

Salo, S., Tahvonen, O., 2002. On equilibrium cycles and normal forests in optimal harvesting of tree vintages. Journal of Environmental Economics and Management 44, 1–22.
Salo, S., Tahvonen, O., 2003. On the economics of forest vintages. Journal of Economics Dynamics and Control 27, 1411–1435.
Samuelson, P.A., 1949. Market mechanisms and maximization: dynamics and linear programming. Rand Corporation, Santa Monica. Reprinted in Stiglitz, J.E. (Ed.), The Collected Scientific Papers of Paul A. Samuelson. vol. II, MIT Press, Cambridge.
Samuelson, P.A., 1971. Maximum principles in analytical economics. Les Prix Nobel en 1970. The Nobel Foundation, Stockholm, 273–288.
Samuelson, P.A., 1976a. Economics of forestry in an evolving society, Economic Inquiry XIV, 466–492.
Samuelson, P.A., 1976b. The periodic turnpike theorem. Nonlinear Analysis, Theory, Methods and Applications 1, 1–13.